

Chapter 2

Postulates and Central Results

Additional Reading: *McQuarrie Ch. 2, Dill and Bromberg Ch. 10*

See also: C. E. Shannon, "A Mathematical Theory of Communication," The Bell System Technical Journal, Vol. 27, pp. 379-423, 623-656 (1948); E.T. Jaynes, "Information Theory and Statistical Mechanics," Phys. Rev. Vol. 106, No. 4, 620-630 (1957)

I. Ensembles with Constraints

In Chapter 1 (II D), two states (α, β) for an assembly of N particles were defined. Practically, every ensemble we deal with is subject to some constraint, such as constant energy, temperature, pressure, or volume.

A. Example: Constant energy ensemble for an assembly of two non-interacting particles in a 3D box.

i. Quantum Particles. The energy for a particle in a cubic 3D box of edge length L is

$$\epsilon_i = \frac{h^2}{8m_i L^2} (n_x^2 + n_y^2 + n_z^2) = \epsilon_0 (n_x^2 + n_y^2 + n_z^2), \quad (2.1)$$

where $\epsilon_0 = h^2/8m_i L^2$. For two particles, the energies sum as

$$E = \sum_{i=1}^2 \epsilon_i. \quad (2.2)$$

With a restriction such as $E = 9\epsilon_0$, 6 states can be enumerated:

state	n_{1x}	n_{1y}	n_{1z}	n_{2x}	n_{2y}	n_{2z}
α	2	1	1	1	1	1
β	1	2	1	1	1	1
γ	1	1	2	1	1	1
δ	1	1	1	2	1	1
η	1	1	1	1	2	1
κ	1	1	1	1	1	2

ii. Classical Particles. The energy for a classical particle in a box is

$$\epsilon_i = \frac{p_{ix}^2 + p_{iy}^2 + p_{iz}^2}{2m} \quad (2.3)$$

There are an infinite number of ways (due to the infinity of real numbers, and therefore infinitely different momenta and positions) for which a two particle assembly can achieve $E = 9\epsilon_0$.

Phase Space. A useful way to depict allowed states is with a phase space plot. The above example (considering momenta and position of two particles in three dimensions - twelve degrees of freedom) cannot easily be depicted in a two-dimensional plot; try to visualize multidimensional extensions of the following phase space plots for the one particle, constant energy ensembles.

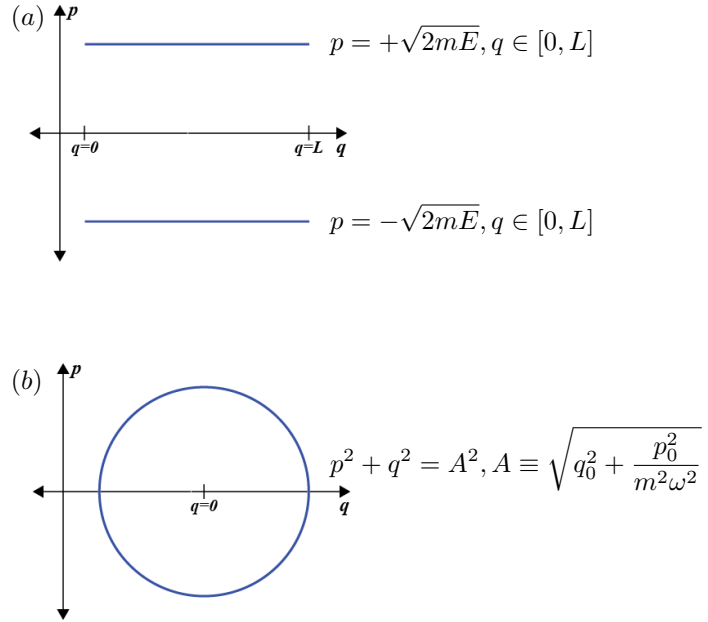


Figure 2.1: Phase space plots for classical (a) particle in a box and (b) harmonic oscillator

B. **Canonical Ensemble.** This ensemble is subject to the constraint that N , V , T are constant, which corresponds to a closed, thermodynamically stable system. In such a system, the Helmholtz free energy $A = E - TS$ is minimized.

II. Derivation of Partition Function for Canonical Ensemble

- A. We know from the idea of ensemble theory that we can replace the time-average of a macroscopic property,

$$\langle f \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau f(t) dt, \quad (2.4)$$

(where $\langle f \rangle$ is an observed macroscopic quantity and f is a microscopic mechanical variable associated with operator F), with the ensemble average,

$$\langle f \rangle = \sum_j^N P_j f_j \quad (2.5)$$

$$\langle f \rangle = \int \dots \int dq^{3N} dp^{3N} P(q^{3N}, p^{3N}) f(q^{3N}, p^{3N}), \quad (2.6)$$

for a quantum or classical system, respectively.

Thus, we begin with a key question: what is the functional form of P_j ?

B. Derivation 1

The Helmholtz free energy can be written in terms of the P_j values with the definition of the ensemble averaged internal energy,

$$E = \sum_j^N P_j E_j \quad (2.7)$$

$$A = E - TS \quad (2.8)$$

$$= \sum_j^N E_j P_j - TS. \quad (2.9)$$

Temperature is a control variable that is constant in this ensemble. What about the entropy S ? For this, we rely on the central assumption of Ludwig Boltzmann (please see CE Shannon and ET Jaynes for more detail about why this must be the form of S),

$$S = -k_B \sum_j^N P_j \ln P_j, \quad (2.10)$$

where Boltzmann's constant, $k_B = R/N_A = 1.38 \cdot 10^{-23} J/K$. Thus,

$$A = \sum_j^N E_j P_j + k_B T \sum_j^N P_j \ln P_j \quad (2.11)$$

$$= \sum_j^N P_j (E_j + k_B T \ln P_j). \quad (2.12)$$

We are searching for the form of P_j that minimizes the Helmholtz free energy. As a stable minimum, if we slightly perturb the $P_j \rightarrow P_j + \delta P_j$, A is unchanged, namely $A \rightarrow A + \delta A, \delta A = 0$.

$$\delta A = 0 = \delta \left[\sum_j^N P_j (E_j + k_B T \ln P_j) \right] \quad (2.13)$$

$$= \sum_j^N (E_j \delta P_j) + k_B T \ln P_j \delta P_j + k_B T P_j \frac{1}{P_j} \delta P_j \quad (2.14)$$

$$= \sum_j^N \delta P_j [E_j + k_B T (\ln P_j + 1)]. \quad (2.15)$$

In addition to minimizing A , another constraint is that the probabilities add to 1, before and after perturbation,

$$\sum_j^N P_j = 1 \quad (2.16)$$

$$\sum_j^N (P_j + \delta P_j) = 1. \quad (2.17)$$

This implies that the probability perturbations sum to 0,

$$\sum_j^N \delta P_j = 0, \quad (2.18)$$

and any one chosen probability fulfills the property,

$$\delta P_1 = - \sum_{j=2}^N \delta P_j. \quad (2.19)$$

Thus,

$$\delta A = 0 = \delta P_1 [E_1 + k_B T (\ln P_1 + 1)] \quad (2.20)$$

$$+ \sum_{j=2}^N \delta P_j [E_j + k_B T (\ln P_j + 1)] \quad (2.21)$$

$$= \sum_{j=2}^N \delta P_j [(E_j - E_1) + k_B T (\ln P_j - \ln P_1)]. \quad (2.22)$$

Since the δP_j terms from $j = 2$ to N are independent for arbitrary δP_j (we can perturb any individual probability in the + or - direction

as we choose),

$$[(E_j - E_1) + k_B T (\ln P_j - \ln P_1)] = 0 \quad (2.23)$$

$$\frac{P_j}{P_1} = \frac{e^{-E_j/k_B T}}{e^{-E_1/k_B T}} \quad (2.24)$$

$$P_j = P_1 e^{E_1/k_B T} e^{-E_j/k_B T}. \quad (2.25)$$

Owing to the normalization condition,

$$1 = \sum_j^N P_j \quad (2.26)$$

$$= P_1 e^{E_1/k_B T} \sum_j^N e^{-E_j/k_B T}, \quad (2.27)$$

we find that

$$P_1 = \frac{e^{-E_1/k_B T}}{\sum_j^N e^{-E_j/k_B T}}, \quad (2.28)$$

or that P_1 or any P_j all have the same form, thus we can simply write,

$$P_j = \frac{e^{-E_j/k_B T}}{\sum_j^N e^{-E_j/k_B T}}. \quad (2.29)$$

This is the **Canonical Distribution** and it gives the probability for the j^{th} distinguishable state in the ensemble. This distribution minimizes A and is thus the equilibrium distribution in the N, V, T ensemble.

The denominator is the **Canonical Partition Function**

$$Q(N, V, T) = \sum_j^N e^{-E_j/k_B T}. \quad (2.30)$$

- C. **Derivation 2** This derivation uses Lagrange multipliers to again find the form of P_j . In this case, we would like to calculate the expectation value of operators, $\langle F \rangle$ under the constraint that the entropy is maximized, which will be true in equilibrium for any closed system.

First, a brief review of Lagrange multipliers.

Lagrange Multipliers in One Page

This method is useful for finding the functional form that maximizes a function of n variables, $f(x_1, x_2, \dots, x_n)$, given a set of m constraints $g_1(x, y, \dots, z) = c_1, g_2(x, y, \dots, z) = c_2, \dots, g_m(x_1, x_2, \dots, x_n) = c_m$, where the set of cs are constants. To do so, we first define a quantity

$$J = f(x_1, x_2, \dots, x_n) - \sum_{j=1}^m \lambda_j [g_j(x_1, x_2, \dots, x_n) - c_j], \quad (2.31)$$

where λ_j is the Lagrange multiplier. We then solve the conditions

$$\frac{\partial J}{\partial x_i} = 0, \frac{\partial J}{\partial \lambda_j} = 0 \quad (2.32)$$

This gives us n equations (for each variable x_i) and m equations (for each constraint or Lagrange multiplier, λ_j) for a total of $n + m$ equations.

- i. **Example:** Find the dimensions of a rectangular solid whose volume, $V(x, y, z) = xyz$, is maximized for a fixed surface area, $A(x, y, z) = (2xy + 2yz + 2xz) = s$, where s is a constant.
- ii. **Solution:**

$$J = V(x, y, z) - \lambda (g(x, y, z) - s) \quad (2.33)$$

$$J = xyz - \lambda (2xy + 2yz + 2xz - s) \quad (2.34)$$

Now we solve for λ ,

$$\frac{\partial J}{\partial x} = \frac{\partial}{\partial x} (xyz - \lambda (2xy + 2yz + 2xz - s)) = 0 \quad (2.35)$$

$$yz - \lambda(2y + 2z) = 0 \quad (2.36)$$

$$\lambda = \frac{yz}{2y + 2z}, \quad (2.37)$$

The second equation can be solved when substituting the above for λ ,

$$\frac{\partial J}{\partial y} = xz - \frac{yz}{2y + 2z} (2x + 2z) = 0 \quad (2.38)$$

$$2xyz + 2xz^2 = 2xyz + 2yz^2 \quad (2.39)$$

$$x = y. \quad (2.40)$$

Plugging this λ into the third equation similarly yields,

$$\frac{\partial J}{\partial z} = xy - \frac{yz}{2y + 2z} (2y + 2x) = 0 \quad (2.41)$$

$$2xy^2 + 2xyz = 2xyz + 2y^2z \quad (2.42)$$

$$x = z. \quad (2.43)$$

Now we know that $x = y = z$. The derivative with respect to λ gives us the original constraint,

$$\frac{\partial J}{\partial \lambda} = -[2(xy + yz + zx) - s] = 0 \quad (2.44)$$

$$6x^2 = s \quad (2.45)$$

$$x = y = z = \sqrt{\frac{s}{6}} \quad (2.46)$$

We would like to find the form of the probabilities that maximize the entropy as a function of the set of P_j ,

$$S(\{P_j\}) = -k_B \sum_j^N P_j \ln P_j, \quad (2.47)$$

given the constraints that

$$\sum_j^N P_j = 1, \sum_j^N P_j E_j = \langle E \rangle \quad (2.48)$$

We first define J such that

$$J = -k_B \sum_j^N P_j \ln P_j - \lambda \left(\sum_j^N P_j - 1 \right) - \beta \left(\sum_j^N P_j E_j - \langle E \rangle \right) \quad (2.49)$$

First we solve

$$\frac{\partial J}{\partial P_j} = -k_B (\ln P_j + 1) - \lambda - \beta E_j = 0 \quad (2.50)$$

$$\ln P_j = -\frac{\lambda + \beta E_j}{k_B} - 1 \quad (2.51)$$

$$P_j = \exp \left[\frac{-\lambda}{k_B} - 1 \right] \exp \left[\frac{-\beta E_j}{k_B} \right] \quad (2.52)$$

From the constraint

$$\frac{\partial J}{\partial \lambda} = \sum_j^N P_j - 1 = 0 \quad (2.53)$$

$$\exp \left[\frac{-\lambda}{k_B} - 1 \right] \sum_j^N \exp \left[\frac{-\beta E_j}{k_B} \right] = 1 \quad (2.54)$$

$$\frac{1}{\exp \left[\frac{\lambda}{k_B} + 1 \right]} \sum_j^N \exp \left[\frac{-\beta E_j}{k_B} \right] = 1 \quad (2.55)$$

we define λ in such a way that

$$\exp \left[\frac{\lambda}{k_B} + 1 \right] = \sum_j^N \exp \left[\frac{-\beta E_j}{k_B} \right] = Q. \quad (2.56)$$

The functional form of λ as the logarithm of the sum is not important to us; we now solve for β given that

$$P_j = \frac{\exp \left(\frac{-\beta E_j}{k_B} \right)}{Q}. \quad (2.57)$$

To do this, we plug P_j back into the definition of S ,

$$S = -k_B \sum_j^N P_j \ln P_j \quad (2.58)$$

$$= -k_B \sum_j^N P_j \left(\frac{-\beta E_j}{k_B} - \ln(Q) \right) \quad (2.59)$$

$$= \beta \sum_j^N P_j E_j + k_B \ln(Q) \quad (2.60)$$

$$= \beta \langle E \rangle + k_B \ln(Q). \quad (2.61)$$

The final step is to use the thermodynamic relation that defines temperature,

$$\frac{dS}{d\langle E \rangle} = \frac{1}{T} \quad (2.62)$$

$$\frac{d}{d\langle E \rangle} (\beta \langle E \rangle + k_B \ln(Q)) = \frac{1}{T} \quad (2.63)$$

$$\beta = \frac{1}{T}. \quad (2.64)$$

Finally we have the Canonical Distribution,

$$P_j = \frac{e^{-E_j/k_B T}}{\sum_j^N e^{-E_j/k_B T}} = \frac{1}{Q} e^{-E_j/k_B T} \quad (2.65)$$

and the Canonical Partition Function,

$$Q(N, V, T) = \sum_j^N e^{-E_j/k_B T}. \quad (2.66)$$

III. Thermodynamic Functions and Macroscopic Properties in Terms of Q

From here on out, we'll define $\beta = 1/k_B T$, which has units of inverse energy. You should memorize $k_B T$ in some of your favorite unit systems:

$$k_B T = 200 \text{ cm}^{-1} = 4.114 \text{ pN} \cdot \text{nm} = 0.0257 \text{ eV} \quad (2.67)$$

$$= 2.479 \text{ kJ/mol} = 0.593 \text{ kcal/mol} \quad (2.68)$$

It also makes differentiation easier. To show this, let's calculate some thermodynamic functions from Q . We'll start with $\langle E \rangle$. Since E_j is the in the argument of the exponential, we'll guess that a partial derivative with respect to β might be useful.

$$Q(N, V, T) = \sum_j e^{-\beta E_j} \quad (2.69)$$

$$\frac{\partial Q}{\partial \beta} = - \sum_j E_j e^{-\beta E_j} = -Q \langle E \rangle \quad (2.70)$$

We can rearrange terms to get

$$\langle E \rangle = -\frac{1}{Q} \frac{\partial Q}{\partial \beta} = -\frac{\partial \ln Q}{\partial \beta}. \quad (2.71)$$

Recall that we can do the following to convert from functions of β back to T ,

$$\frac{\partial \beta}{\partial T} = \frac{-1}{k_B T^2} \quad (2.72)$$

$$1 = -k_B T^2 \frac{\partial \beta}{\partial T} \quad (2.73)$$

$$\langle E \rangle = -\frac{\partial \ln Q}{\partial \beta} \left(-k_B T^2 \frac{\partial \beta}{\partial T} \right) \quad (2.74)$$

$$\langle E \rangle = k_B T^2 \frac{\partial \ln Q}{\partial T} \quad (2.75)$$

In addition to $\langle E \rangle$, S can be written in terms of Q ,

$$S = -k_B \sum_j p_j \ln(p_j) \quad (2.76)$$

$$= -k_B \sum_j p_j \ln \left(\frac{e^{-\beta E_j}}{Q} \right) \quad (2.77)$$

$$= -k_B \sum_j p_j \left(-\frac{E_j}{k_B T} - \ln Q \right) \quad (2.78)$$

$$= \frac{\sum_j p_j E_j}{T} + k_B \ln Q \quad (2.79)$$

$$= \frac{\langle E \rangle}{T} + k_B \ln Q \quad (2.80)$$

$$= k_B T \frac{\partial \ln Q}{\partial T} + k_B \ln Q. \quad (2.81)$$

With expressions for $\langle E \rangle$ and S , we can calculate the Helmholtz free energy,

$$A = \langle E \rangle - TS \quad (2.82)$$

$$= \langle E \rangle - T \left(k_B \ln Q + \frac{\langle E \rangle}{T} \right) \quad (2.83)$$

$$= -k_B T \ln Q, \quad (2.84)$$

which is a surprisingly simple result!

Since the Helmholtz free energy is the form of energy that is minimized in the canonical ensemble, it enables the calculation of many other quantities,

$$p = - \left(\frac{\partial A}{\partial V} \right)_{T,N} = k_B T \left(\frac{\partial \ln Q}{\partial V} \right)_{T,N} \quad (2.85)$$

$$H = \langle E \rangle + pV = \langle E \rangle + k_B T V \left(\frac{\partial \ln Q}{\partial V} \right)_{T,N} \quad (2.86)$$

$$G = A + pV = k_B T^2 \left(\frac{\partial \ln Q}{\partial T} \right)_{T,N} + k_B T V \left(\frac{\partial \ln Q}{\partial V} \right)_{T,N} \quad (2.87)$$

$$\mu = \left(\frac{\partial A}{\partial N} \right)_{T,V} = -k_B T \left(\frac{\partial \ln Q}{\partial N} \right)_{T,V} \quad (2.88)$$

IV. Evaluating Q for Simple Models

When the form of the energy involves only sums or pair-wise, nearest neighbor interactions, we can solve the model exactly. The most famous of this class of models is the Ising Model, first solved in 1D by Ernst Ising.

In this model, there are N spins, each situated at a lattice site, and each spin may be up or down.



Each $\boxed{\uparrow}$ adds J to the total energy of the system and each $\boxed{\downarrow}$ stabilizes it by J . If neighboring spins are aligned, such as $\boxed{\uparrow\uparrow}$ or $\boxed{\downarrow\downarrow}$, there is an additional stabilization of K . If they are not aligned, such as $\boxed{\uparrow\downarrow}$ or $\boxed{\downarrow\uparrow}$, there is an energy penalty of K .

If we represent an $\boxed{\uparrow}$ with $s_i = +1$ and $\boxed{\downarrow}$ with $s_i = -1$, then the energy function for the system, the Hamiltonian, is

$$H = -J \sum_i s_i - K \sum_{|i-j|=1} s_i s_j, \quad (2.89)$$

namely a sum over all sites and a sum over all nearest neighbors.

A. Case 1: Three spins with $J = 0$

Let's work out the partition function and the average energy for three spins, where's its feasible to manually enumerate all the states. When $J = 0$,

$$\begin{array}{|c|c|c|} \hline \uparrow & \uparrow & \uparrow \\ \hline \end{array} \quad E = 2K \quad (2.90)$$

$$\begin{array}{|c|c|c|} \hline \downarrow & \uparrow & \uparrow \\ \hline \end{array} \quad E = 0 \quad (2.91)$$

$$\begin{array}{|c|c|c|} \hline \uparrow & \downarrow & \uparrow \\ \hline \end{array} \quad E = -2K \quad (2.92)$$

$$\begin{array}{|c|c|c|} \hline \uparrow & \uparrow & \downarrow \\ \hline \end{array} \quad E = 0 \quad (2.93)$$

$$\begin{array}{|c|c|c|} \hline \uparrow & \downarrow & \downarrow \\ \hline \end{array} \quad E = 0 \quad (2.94)$$

$$\begin{array}{|c|c|c|} \hline \downarrow & \uparrow & \downarrow \\ \hline \end{array} \quad E = -2K \quad (2.95)$$

$$\begin{array}{|c|c|c|} \hline \downarrow & \downarrow & \uparrow \\ \hline \end{array} \quad E = 0 \quad (2.96)$$

$$\begin{array}{|c|c|c|} \hline \downarrow & \downarrow & \downarrow \\ \hline \end{array} \quad E = 2K \quad (2.97)$$

The partition function is then,

$$Q = \sum_j e^{-\beta E_j} \quad (2.98)$$

$$= 2e^{2\beta K} + 4e^0 + 2e^{-2\beta K}. \quad (2.99)$$

Recalling that $\cosh(x) = \frac{1}{2}(e^x + e^{-x})$,

$$Q = 2^3 \cosh^2(\beta K). \quad (2.100)$$

then the average energy is,

$$\langle E \rangle = -\frac{1}{Q} \frac{\partial Q}{\partial \beta} \quad (2.101)$$

$$= -\frac{1}{2^3 \cosh^2(\beta K)} \cdot 2^4 K \cosh(\beta K) \sinh(\beta K) \quad (2.102)$$

$$= -2K \frac{\sinh(\beta K)}{\cosh(\beta K)} \quad (2.103)$$

$$= -2K \tanh(\beta K). \quad (2.104)$$

We'd like to use this result to visualize the states that are populated as we change external variables. In this case, we can change only T . As $T \rightarrow 0$ or $\beta \rightarrow \infty$,

$$\lim_{\beta \rightarrow \infty} \langle E \rangle = -2K \lim_{\beta \rightarrow \infty} \tanh(\beta K) \quad (2.105)$$

$$= -2K \lim_{\beta \rightarrow \infty} \frac{e^{\beta K} - e^{-\beta K}}{e^{\beta K} + e^{-\beta K}} \quad (2.106)$$

$$= -2K, \quad (2.107)$$

which tells us that as the system is cooled, only the lowest energy states, $\boxed{\uparrow\downarrow\uparrow}$ and $\boxed{\downarrow\uparrow\downarrow}$ are occupied. As $T \rightarrow \infty$ or $\beta \rightarrow 0$,

$$\lim_{\beta \rightarrow 0} \langle E \rangle = -2K \lim_{\beta \rightarrow 0} \tanh(\beta K) \quad (2.108)$$

$$= -2K \frac{e^0 - e^0}{e^0 + e^0} \quad (2.109)$$

$$= 0, \quad (2.110)$$

which fits our expectation that as the system is heated, the difference between energy levels $2K$ becomes negligible and all states are equally probable. (If all the energies 2.90 - 2.97 are equally probable, then the average energy will be 0.)

B. Case 2: N spins with $K = 0$

In this case, the Hamiltonian will simply be,

$$H = -J \sum_i s_i, \quad (2.111)$$

and the partition function is then,

$$Q = \sum_{\{s_i\}} e^{\beta J \sum_i s_i}. \quad (2.112)$$

The sum $\sum_{\{s_i\}}$ is a sum over all states, such as $\{-1, -1, 1, 1, -1, 1, 1, -1, \dots\}$. How can we simplify this expression?

The key to simplifying this expression is to realize that the sum in the exponential can be broken up into a product,

$$e^{x_1+x_2+\dots} = e^{x_1} \cdot e^{x_2} \dots \quad (2.113)$$

The partition function can then be written as,

$$Q = \sum_{\{s_i\}} \prod_i e^{\beta J s_i}. \quad (2.114)$$

Since each spin is independent, we can rearrange terms as

$$Q = \prod_i \sum_{s_i=\{-1,+1\}} e^{\beta J s_i} \quad (2.115)$$

or a product of N terms, one for each site which can be $s_i = +1$ or $s_i = -1$. This can be simplified further using the trigonometric identities,

$$Q = \left(\sum_{s_i=\{-1,+1\}} e^{\beta J s_i} \right)^N \quad (2.116)$$

$$= (e^{\beta J \cdot 1} + e^{-\beta J \cdot -1})^N \quad (2.117)$$

$$= 2^N \cosh^N(\beta J) \quad (2.118)$$

If we look at the average energy and its temperature limits,

$$\langle E \rangle = -\frac{1}{Q} \frac{\partial Q}{\partial \beta} \quad (2.119)$$

$$= -\frac{1}{2^N \cosh^N(\beta J)} 2^N N J \sinh^{N-1}(\beta J) \cosh(\beta J) \quad (2.120)$$

$$= -N J \frac{\sinh(\beta J)}{\cosh(\beta J)} \quad (2.121)$$

$$= -N J \tanh(\beta J), \quad (2.122)$$

which $\rightarrow 0$, as $\beta \rightarrow 0$ and $\rightarrow -NJ$, as $\beta \rightarrow \infty$ using the limits worked out as in 2.105-2.110. Again, we note that as temperature increases, $\boxed{\uparrow}$ and $\boxed{\downarrow}$ are equally likely at each site. The average energy is $\frac{1}{2} [J + (-J)] = 0$ for each site and the average energy of the whole system is $0 \cdot N = 0$. As the system is cooled, only the lowest energy state is occupied (each spin as $\boxed{\uparrow}$ and contributing $-J$ per site or $-NJ$ total.)

V. Microcanonical Ensemble A somewhat simpler collection of assembly states than the Canonical Ensemble (fixed N,V,T) is the Microcanonical Ensemble with fixed particle number, volume, and energy (N,V,E). We can construct a whole set of different ensembles in order to match experimental constraints or to simplify the math. We'll find that in the macroscopic limit, most problems will yield the same result.

A. Microcanonical Partition Function

We employ a similar procedure to derive the Microcanonical Partition Function as what we used to derive the Canonical Partition Function; that is, we frame it as a problem suitable for Lagrange Multipliers. What is the functional form of the probabilities of the assembly states, p_j given that the entropy, $S = -k_B \sum_j^\Omega p_j \ln(p_j)$ is maximized subject to the normalization constraint, $\sum_j^\Omega p_j = 1$, where Ω is the number of states with energy E.

As before, we construct the function, J , and set its derivatives equal to zero,

$$J = -k_B \sum_j^\Omega p_j \ln(p_j) - \lambda \left[\sum_j^\Omega p_j - 1 \right] \quad (2.123)$$

$$\frac{\partial J}{\partial p_j} = -k_B [\ln(p_j) + 1] - \lambda = 0 \quad (2.124)$$

$$\ln(p_j) = -\frac{\lambda}{k_B} - 1 \quad (2.125)$$

$$p_j = e^{-(1+\lambda/k_B)}. \quad (2.126)$$

The derivative with respect to λ gives us the normalization condition,

$$\frac{\partial J}{\partial \lambda} = \sum_j^{\Omega} p_j - 1 = 0 \quad (2.127)$$

$$1 = \sum_j^{\Omega} p_j \quad (2.128)$$

$$= \sum_j^{\Omega} e^{-(1+\lambda/k_B)} \quad (2.129)$$

$$= \Omega e^{-(1+\lambda/k_B)}. \quad (2.130)$$

We define λ such that

$$\frac{1}{\Omega} = e^{-(1+\lambda/k_B)}, \quad (2.131)$$

and a more explicit expression for λ is not necessary. Comparing 2.126 and 2.131 gives us the desired result,

$$p_j = \frac{1}{\Omega}. \quad (2.132)$$

Apparently, any state is equally likely!

For mathematical convenience, we define $\Omega(N, V, E)$ as the number of microscopic states with fixed N , V , and energy between E and $E - \delta E$. This definition allows us to differentiate Ω without getting wild fluctuations.

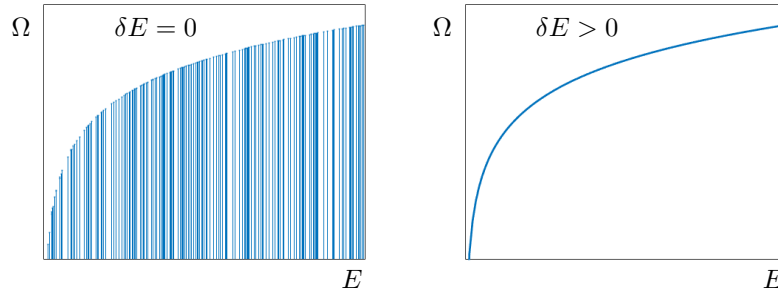


Figure 2.2: (Left) Representation of the number of microscopic states as a function of the discrete energy levels of molecular systems. (Right) By defining Ω to include a small "smearing factor" δE , we get a smooth function with continuous derivatives.

An important related quantity is $\bar{\Omega}(N, V, E)$, which is defined such that $\bar{\Omega}(N, V, E)dE$ is number of states with energy between E and dE . Note that $\bar{\Omega}(N, V, E)$ has units of 1/Energy.

B. Microcanonical Entropy

Let's calculate the entropy for the form of the p_j in the Microcanonical Ensemble,

$$S = -k_B \sum_j p_j \ln(p_j) \quad (2.133)$$

$$= -k_B \sum_j \frac{1}{\Omega} \ln\left(\frac{1}{\Omega}\right) \quad (2.134)$$

$$= -k_B \Omega \frac{1}{\Omega} \ln\left(\frac{1}{\Omega}\right) \quad (2.135)$$

$$= k_B \ln(\Omega). \quad (2.136)$$

This simple and powerful relation showing a logarithmic connection between entropy and probability is very useful to keep in mind. Although we've derived it for fixed-energy systems, the results of the next section shows that it can also be of broader use.

C. Principle of Ensemble Equivalence

A remarkable and surprising result connects the partition functions for the Canonical and Microcanonical Ensembles:

$$Q(N, V, T) = \sum_j e^{-\beta E_j} (\text{sum over states}) \quad (2.137)$$

$$= \sum_l \Omega(E_l) e^{-\beta E_l} (\text{sum over levels}) \quad (2.138)$$

$$\rightarrow \int_0^\infty dE \bar{\Omega}(N, V, E) e^{-\beta E}. \quad (2.139)$$

The last line is appropriate in the macroscopic limit as the energy levels become spaced more and more close together. The last line can also be identified with a Laplace Transform,

$$\int_0^\infty dE \bar{\Omega}(N, V, E) e^{-\beta E} = \mathcal{L} [\bar{\Omega}(N, V, E)]. \quad (2.140)$$

Thus the canonical partition function, $Q(N, V, T)$ is the Laplace transform of the microcanonical $\bar{\Omega}(N, V, E)$ in the macroscopic limit. The Laplace transform is a unique transformation, and thus these two functions contain identical information. How can this be, given that different variables are fixed in each function? A fixed temperature implies that a distribution of different energy states is populated. Also, an ensemble of fixed energy cannot be in thermal equilibrium and thus has no defined temperature (recall $\frac{1}{T} = \frac{dS}{dE} \big|_{V, T}$, or the distribution of energies given by the Maxwell-Boltzmann distribution for an ideal gas at fixed temperature).

We can resolve this apparent paradox by quantifying the expectation value for energy fluctuations in relation to the average energy in the canonical ensemble, which is:

$$\frac{\sqrt{\langle (\delta E)^2 \rangle}}{\langle E \rangle}, \quad (2.141)$$

where

$$\delta E \equiv E - \langle E \rangle. \quad (2.142)$$

We can simplify the numerator in 2.141 using,

$$\langle (\delta E)^2 \rangle = \langle (E - \langle E \rangle)^2 \rangle \quad (2.143)$$

$$= \langle E^2 \rangle - 2 \langle E \rangle \langle E \rangle + \langle E \rangle^2 \quad (2.144)$$

$$= \langle E^2 \rangle - 2 \langle E \rangle \langle E \rangle + \langle E \rangle^2 \quad (2.145)$$

$$= \langle E^2 \rangle - \langle E \rangle^2. \quad (2.146)$$

It was derived in 2.71 that

$$\langle E \rangle = -\frac{1}{Q} \frac{\partial Q}{\partial \beta}. \quad (2.147)$$

We can also derive a relation between $\langle E^2 \rangle$ and a derivative of Q ,

$$\frac{\partial^2 Q}{\partial \beta^2} = \sum_j E_j^2 e^{-\beta E_j} = \sum_j E_j^2 p_j Q = \langle E^2 \rangle Q, \quad (2.148)$$

and thus

$$\langle E^2 \rangle = \frac{1}{Q} \frac{\partial^2 Q}{\partial \beta^2}. \quad (2.149)$$

Substituting this relation for $\langle (\delta E)^2 \rangle$, we find

$$\langle E^2 \rangle - \langle E \rangle^2 = \frac{1}{Q} \frac{\partial^2 Q}{\partial \beta^2} - \langle E \rangle^2 \quad (2.150)$$

$$= \frac{1}{Q} \frac{\partial}{\partial \beta} \frac{\partial Q}{\partial \beta} - \langle E \rangle^2 \quad (2.151)$$

$$= \frac{1}{Q} \frac{\partial}{\partial \beta} (-Q \langle E \rangle) - \langle E \rangle^2 \quad (2.152)$$

$$= \frac{1}{Q} \left(-\frac{\partial Q}{\partial \beta} \langle E \rangle - Q \frac{\partial \langle E \rangle}{\partial \beta} \right) - \langle E \rangle^2 \quad (2.153)$$

$$= \frac{1}{Q} \left(Q \langle E \rangle \langle E \rangle - Q \frac{\partial \langle E \rangle}{\partial \beta} \right) - \langle E \rangle^2 \quad (2.154)$$

$$= -\frac{\partial \langle E \rangle}{\partial \beta} = kT^2 \frac{\partial \langle E \rangle}{\partial T} \quad (2.155)$$

(For the last line, see 2.72-2.75 for how to convert between derivatives with respect to β and T .)

Here if we recall a bit of thermodynamics (that $C_V \equiv \frac{\partial \langle E \rangle}{\partial T}$),

$$\frac{\sqrt{\langle (\delta E)^2 \rangle}}{\langle E \rangle} = \frac{\sqrt{k_B T^2 C_V}}{\langle E \rangle}. \quad (2.156)$$

To make a more solid connection, we can use a relation that we will derive in the next section,

$$C_V \propto N k_B, \langle E \rangle \propto N k_B T, \quad (2.157)$$

and so

$$\frac{\sqrt{\langle (\delta E)^2 \rangle}}{\langle E \rangle} = \frac{\sqrt{k_B T^2 N k_B}}{N k_B T} \quad (2.158)$$

$$= \frac{\sqrt{N}}{N} \quad (2.159)$$

$$= \frac{1}{\sqrt{N}}. \quad (2.160)$$

This last relation explains the apparent paradox in the equivalence of the canonical and microcanonical ensembles. For $N = 10^{22}$, $\frac{1}{\sqrt{N}} \approx 10^{-11}$. While in the canonical ensemble the energy is allowed to fluctuate, as $N \rightarrow \infty$, these energy fluctuations become negligible and observables are identical to those calculated from the microcanonical ensemble.

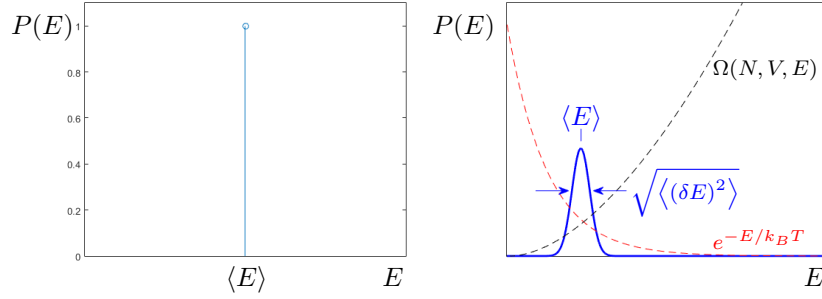


Figure 2.3: (Left) In the microcanonical ensemble, the energy is fixed and the probability is unity for finding that energy and precisely zero for all other energies. (Right) In the canonical ensemble, the energy can fluctuate; these fluctuations are limited by the decreasing density of states at lower energies and the decreasing Boltzmann factor at high energies. As the system size becomes larger, the spread in probable energies becomes smaller and approaches an infinitely narrow distribution centered at $\langle E \rangle$, as in the microcanonical ensemble.