

Chapter 3

Entropy

3.1 Multiplicity

The 16 possible outcomes of flipping a fair coin 4 times are shown in Table 3.1. Let's define the number of heads obtained from these coin flips to be a *macrostate* of this system. The *microstates* associated with each macrostate are the specific combinations of heads and tails that correspond to the macrostate; these microstates are also shown in Table 3.1. We define the number of microstates associated with each macrostate as the *multiplicity* of the macrostate; the variable Ω is used to denote multiplicity. For example, the multiplicity of obtaining n heads out of a total of N flips is¹

$$\Omega_N(n) = \frac{N!}{n!(N-n)!} \quad (3.1)$$

The macrostate with the highest multiplicity is most likely to be observed as it is associated the largest number of microstates (assuming that all microstates are equally probable). For example, there are 6 ways to obtain 2 heads when flipping a fair coin 4 times, but only 1 way to obtain 4 heads. We would therefore expect that it is more likely that we would obtain 2 heads than 4 heads when flipping a fair coin 4 times. Indeed, if we assume an equal probability of obtaining heads and obtaining tails for each flip (a probability of 0.5 for each outcome) then the probability of obtaining any individual microstate of four coin flips would be

$$p = \left(\frac{1}{2}\right)^4 \rightarrow p = \frac{1}{16}$$

Since there is only one microstate associated with obtaining 4 heads, the probability of obtaining 4 heads is thus $\frac{1}{16}$. However, since there are 6 ways of obtaining 2 heads, the probability of obtaining 2 heads is

¹This is also the equation for the number of combinations of selecting n objects out of a total of N possible objects.

Macrostate (Number of Heads)	Microstates	Multiplicity (Ω) (Number of Microstates)	Probability
4	HHHH	1	$\frac{1}{16}$
3	THHH HTHH HHTH HHHT	4	$\frac{1}{4}$
2	TTHH THTH THHT HTTH HTHT HHTT	6	$\frac{3}{8}$
3	HTTT THTT TTHT TTTH	4	$\frac{1}{4}$
4	HHHH	1	$\frac{1}{16}$

Table 3.1: The 16 possible outcomes of flipping a fair coin four times. We can define the number of heads obtained as the macrostate of this process/measurement. The microstates corresponding to each macrostate, the associated multiplicity, and probability of obtaining the macrostate are also shown.

$$6 \cdot \frac{1}{16} = \frac{3}{8}$$

As expected, the probability of obtaining 4 heads is larger than the probability of obtaining 1 head.

The probability of finding a thermodynamic system in its n th accessible macrostate is equal to the multiplicity of that macrostate divided by the sum of all possible multiplicities.

$$p_n = \frac{\Omega_n}{\sum_n \Omega_n}$$

For example, the probability of observing the “2 heads” macrostate is equal to the ratio of the multiplicity of that macrostate (6) to the sum of the multiplicities of all macrostates ($1+4+6+4+1 = 16$), which equals $\frac{6}{16} = \frac{3}{8}$ as calculated earlier.

Generalizing Further

The probability of finding a system in a particular macrostate can be derived from the fundamental postulate of statistical mechanics.

An isolated system in equilibrium is equally likely to be in any of its accessible states, each of which is defined by a particular configuration of the system’s elements.

Therefore the probability of finding an isolated system in equilibrium in a particular macrostate is directly proportional to the multiplicity of that macrostate; that is, it is directly proportional to the number of microstates associated with the macrostate.

Going further, we say that when in equilibrium, systems *will be* in the macrostate corresponding to the highest multiplicity. In effect, we are saying that we expect to find systems in configurations corresponding to the highest probability. This is not necessarily profound, but is nevertheless powerful. Indeed, if we argue that systems in equilibrium are most likely to be found in their most probable state, we can subsequently argue that systems will move spontaneously toward their post probable state in order to achieve equilibrium. *All thermodynamic systems will spontaneously move toward the macrostate with the highest multiplicity.*

3.1.1 Different Statistical Systems

Let's consider a system that contains N particles distributed across degenerate quantum states. The i^{th} energy state has a degeneracy g_i and contains a number of particles n_i . It follows that the total number of particles in the system is determined by

$$\sum_i n_i = N \quad (3.2)$$

Boltzmann Statistics

The number of ways of selecting n_1 particles out of a total of N particles for the first energy state is

$$\frac{N!}{n_1! (N - n_1)!}$$

If the degeneracy of this energy state is g_1 , there are g_1 possible places to place each particle in this energy state, which results in a total of $g_1^{n_1}$ total possibilities. Thus, if the particles are distinguishable, the number of ways to put n_1 particles (out of a total of N particles) into the first quantum state is with degeneracy g_1 (*i.e.*, the multiplicity) is

$$\Omega_1 = g_1^{n_1} \left(\frac{N!}{n_1! (N - n_1)!} \right)$$

For the second quantum state, we are selecting n_2 particles out of a total of $(N - n_1)$ particles to distribute about the g_2 degenerate options. The multiplicity for this state is thus

$$\Omega_2 = g_2^{n_2} \left(\frac{(N - n_1)!}{n_2! (N - n_1 - n_2)!} \right)$$

The total multiplicity for the entire system of states is thus

$$\begin{aligned}
\Omega &= \left[\frac{N!g_1^{n_1}}{n_1!(N-n_1)!} \right] \left[\frac{(N-n_1)!g_2^{n_2}}{n_2!(N-n_1-n_2)!} \right] \left[\frac{(N-n_1-n_2)!g_3^{n_3}}{n_3!(N-n_1-n_2-n_3)!} \right] \cdots \\
\Omega &= \frac{N!g_1^{n_1}g_2^{n_2}g_3^{n_3}\cdots}{n_1!n_2!n_3!\cdots} \\
\Omega_B &= N! \prod_i \frac{g_i^{n_i}}{n_i!} \tag{3.3}
\end{aligned}$$

The subscript B denotes that this distribution is referred to as the Boltzmann distribution.

Fermi-Dirac distribution

Let's now assume that the particles are identical and indistinguishable. Let's also assume that the Pauli Exclusion Principle applies so that no quantum state can accept more than one particle. In other words, $n_i < g_i$ for all i . In this case, n_i of the g_i states are filled with one particle and $g_i - n_i$ states are unoccupied. Substitution into Equation 3.1 allows us to determine the multiplicity for the i^{th} energy state of this distribution.

$$\Omega_N(n) = \frac{N!}{n!(N-n)!} \xrightarrow[N=g_i]{n=n_1} \Omega_i = \frac{g_i!}{n_i!(g_i - n_i)!}$$

The total multiplicity is the product of multiplicity of all energy states

$$\Omega_{FD} = \prod_i \frac{g_i!}{n_i!(g_i - n_i)!} \tag{3.4}$$

The subscript FD denotes that this distribution is referred to as the Fermi-Dirac distribution.

Bose-Einstein distribution

Let's now consider the particles to be identical and indistinguishable, but not subject to the Pauli Exclusion Principle. In this case, there are n_i particles in the i^{th} energy state with no restrictions on the number of particles in each g_i quantum state. The multiplicity of the i^{th} energy state is

$$\Omega_N(n) = \frac{N!}{n!(N-n)!} \xrightarrow[N=g_i+n_i-1]{n=n_1} \Omega_i = \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}$$

The total multiplicity is the product over all possible energy states

$$\Omega_{BE} = \prod_i \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!} \tag{3.5}$$

The subscript BE denotes that this distribution is referred to as the Bose-Einstein distribution.

Maxwell-Boltzmann distribution

The Maxwell-Boltzmann distribution is the classical limit of the Fermi-Dirac distribution and the Bose-Einstein distribution. In this limit, very few states are occupied and thus it is unlikely that more than one particle will be in the same state. Thus, we can approximate Equation 3.4 as

$$\Omega_{FD} = \prod_i \frac{g_i (g_i - 1) (g_i - 2) \dots (g_i - n_i + 1) (g_i - n_i)!}{n_i! (g_i - n_i)!}$$

$$\Omega_{FD} = \prod_i \frac{g_i (g_i - 1) (g_i - 2) \dots (g_i - n_i + 1)}{n_i!} \approx \prod_i \frac{g_i^{n_i}}{n_i!}$$

Similarly, for Equation 3.5, we have

$$\Omega_{BE} = \prod_i \frac{(n_i + g_i - 1) (n_i + g_i - 2) \dots (n_i + g_i - n_i) (g_i - 1)!}{n_i! (g_i - 1)!}$$

$$\Omega_{BE} = \prod_i \frac{(n_i + g_i - 1) (n_i + g_i - 2) \dots (n_i + g_i - n_i)}{n_i!} \approx \prod_i \frac{g_i^{n_i}}{n_i!}$$

$$\Omega_{MB} = \prod_i \frac{g_i^{n_i}}{n_i!} \quad (3.6)$$

The Maxwell-Boltzmann distribution is thus the Boltzmann distribution, but with indistinguishable particles. Because of this, the multiplicity of the Maxwell-Boltzmann distribution differs from the multiplicity of the Boltzmann distribution by the $N!$ term.

$$\Omega_{MB} = \frac{\Omega_B}{N!}$$

3.2 Entropy

Now suppose that we have two systems, A and B, in contact with one another. The probability of finding the combined system - System A and System B - in a particular state is equal to the product of the probabilities of finding the two systems in the corresponding states.

$$p_{total} = p_A p_B$$

The variables p_A and p_B in this expression denote the probability that System A and System B, respectively, will be in their specified states. It follows that the total multiplicity of the combined system is the product of the multiplicity of these two systems.

$$\Omega_{total} = \Omega_A \Omega_B$$

At equilibrium, the combined system will be found in its most probable macrostate, which would correspond to a maximum of Ω_{total} . A condition of this maximum is

$$d\Omega_{total} = (d\Omega_A)\Omega_B + \Omega_A(d\Omega_B) = 0$$

Let's now divide this equation by Ω_{total} .

$$\begin{aligned} \frac{d\Omega_{total}}{\Omega_{total}} &= \frac{(d\Omega_A)\Omega_B}{\Omega_{total}} + \frac{\Omega_A(d\Omega_B)}{\Omega_{total}} = 0 \\ \Omega_{total} = \Omega_A\Omega_B \quad \rightarrow \quad \frac{d\Omega_{total}}{\Omega_{total}} &= \frac{d\Omega_A}{\Omega_A} + \frac{d\Omega_B}{\Omega_B} = 0 \end{aligned}$$

Thus,

$$d(\ln\Omega_{total}) = d(\ln\Omega_A) + d(\ln\Omega_B) = 0$$

We refer to the natural logarithm of the multiplicity as the *entropy*.

$$\mathcal{S} = \ln\Omega \tag{3.7}$$

So our condition for equilibrium is that the entropy of the system is an extremum of the system.

$$d\mathcal{S}_{total} = d\mathcal{S}_A + d\mathcal{S}_B = 0$$

We also see that entropy is additive; $\mathcal{S}_{total} = \mathcal{S}_A + \mathcal{S}_B$.

Boltzmann Statistics

The entropy for a Boltzmann distribution can be determined from Equation 3.3 and Equation 3.7.

$$\ln\Omega_B = \ln N! + \sum_i n_i \ln g_i - \ln n_i!$$

The factorials in this equation can be simplified using Stirling's approximation², $\ln N! = \ln \left(\frac{N}{e}\right)^N = N \ln N - N$.

$$\begin{aligned} \mathcal{S}_B &= \ln \left(\frac{N}{e}\right)^N + \sum_i n_i \ln g_i - n_i \ln n_i + n_i \\ \mathcal{S}_B &= \ln \left(\frac{N}{e}\right)^N + \sum_i n_i + \sum_i n_i \ln \left(\frac{g_i}{n_i}\right) \end{aligned} \tag{3.8}$$

²A more accurate expression for Stirling's approximation is $n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$, which corresponds to $\ln(n!) \approx n \ln(n) - n + \frac{1}{2} \ln(2\pi n)$. The error introduced by truncating the $\frac{1}{2} \ln(2\pi n)$ becomes increasing insignificant when $n \geq 30$.

The summation over n_i in this expression is the number of particles in the system (Equation 3.2).

$$\mathcal{S}_B = \ln \left(\frac{N}{e} \right)^N + N - \sum_i n_i \ln \left(\frac{n_i}{g_i} \right)$$

Fermi-Dirac Statistics

Similarly, the equation for the entropy of a Fermi-Dirac distribution can be determined from Equation 3.4 and Equation 3.7.

$$\ln \Omega_{FD} = \sum_i \ln g_i! - \ln n_i! - \ln (g_i - n_i)!$$

As before, we can simplify this expression using Stirling's approximation.

$$\begin{aligned} \ln \Omega_{FD} &= \sum_i g_i \ln g_i - g_i - n_i \ln n_i + n_i - (g_i - n_i) \ln (g_i - n_i) + g_i - n_i \\ \mathcal{S}_{FD} &= \sum_i g_i \ln g_i - n_i \ln n_i - (g_i - n_i) \ln (g_i - n_i) \end{aligned} \quad (3.9)$$

Bose-Einstein Statistics

As expected, the equation for the entropy of a Bose-Einstein distribution can be determined from Equation 3.5 and Equation 3.7.

$$\ln \Omega_{BE} = \sum_i \ln (n_i + g_i - 1)! - \ln n_i! - \ln (g_i - 1)!$$

Unsurprisingly, we can simplify this expression using Stirling's approximation.

$$\begin{aligned} \ln \Omega_{BE} &= \sum_i (n_i + g_i - 1) \ln (n_i + g_i - 1) - (n_i + g_i - 1) - \\ &\quad n_i \ln n_i + n_i - (g_i - 1) \ln (g_i - 1) + (g_i - 1) \end{aligned}$$

$$\mathcal{S}_{BE} = \sum_i (n_i + g_i - 1) \ln (n_i + g_i - 1) - n_i \ln n_i - (g_i - 1) \ln (g_i - 1) \quad (3.10)$$

Maxwell-Boltzmann Statistics

Finally, the equation for the entropy of a Maxwell-Boltzmann distribution can be determined from Equation 3.6 and Equation 3.7.

$$\ln \Omega_{MB} = \sum_i n_i \ln g_i - \ln n_i!$$

The final summation in this equation can be simplified using Stirling's approximation.

$$\begin{aligned} \mathcal{S}_{MB} &= \sum_i n_i \ln g_i - n_i \ln n_i + n_i \\ \mathcal{S}_{MB} &= \sum_i n_i - \sum_i n_i \ln \left(\frac{n_i}{g_i} \right) \end{aligned} \quad (3.11)$$

The first summation in this expression is the number of particles in the system (Equation 3.2).

$$S_{MB} = N - \sum_i n_i \ln \left(\frac{n_i}{g_i} \right)$$

When we compare Equation 3.8 and Equation 3.11, we see that the entropy of a Boltzmann distribution is larger than the entropy of a Maxwell-Boltzmann distribution. This reflects the fact that particles in the Boltzmann distribution are distinguishable whereas particles in the Maxwell-Boltzmann distribution are indistinguishable, and thus that the multiplicity of any state in a Maxwell-Boltzmann distribution is lower than the multiplicity for the corresponding state in a Boltzmann distribution.

3.3 Canonical Distribution

If the energy of the i^{th} energy state is ϵ_i , the mean energy \bar{E} of our system of N particles is determined from the number of particles n_i in the i^{th} energy state using the following equation:

$$\sum_i n_i \epsilon_i = \bar{E} \quad (3.12)$$

Let's determine the expression of n_i that maximizes the entropy of our statistical distributions subject to the constraints in Equation 3.2 and Equation 3.12 using Lagrange multipliers.

Boltzmann Statistics

Maximizing Equation 3.8 yields

$$\begin{aligned} \frac{\partial}{\partial n_i} \left[\ln \left(\frac{N}{e} \right)^N + \sum_i n_i - \sum_i n_i \ln \left(\frac{n_i}{g_i} \right) \right] - \\ \alpha \frac{\partial}{\partial n_i} \left[\sum_i n_i - N \right] - \beta \frac{\partial}{\partial n_i} \left[\sum_i n_i \epsilon_i - \bar{E} \right] = 0 \end{aligned}$$

When taking the derivatives, only terms including n_i are nonzero. This leaves us with

$$1 - \ln\left(\frac{n_i}{g_i}\right) - \frac{n_i}{n_i} - \alpha - \beta\epsilon_i = 0 \quad \rightarrow \quad \ln\left(\frac{n_i}{g_i}\right) = -\alpha - \beta\epsilon_i$$

$$\frac{n_i}{g_i} = e^{-(\alpha + \beta\epsilon_i)} \quad (3.13)$$

Substitution of Equation 3.13 into Equation 3.2 then yields

$$N = \sum_i g_i e^{-(\alpha + \beta\epsilon_i)} \quad \rightarrow \quad N = e^{-\alpha} \sum_i g_i e^{-\beta\epsilon_i}$$

$$e^{-\alpha} = \frac{N}{\sum_i g_i e^{-\beta\epsilon_i}}$$

$$\frac{n_i}{g_i} = \frac{N e^{-\beta\epsilon_i}}{\sum_i g_i e^{-\beta\epsilon_i}} \quad (3.14)$$

We thus recognize that the probability p_i that the i^{th} quantum state is occupied is

$$n_i = p_i N \quad \rightarrow \quad p_i = \frac{g_i e^{-\beta\epsilon_i}}{\sum_i g_i e^{-\beta\epsilon_i}} \quad (3.15)$$

The summation over states in the denominator of Equation 3.14 is referred to as the partition function³ and is denoted by Z .

$$Z = \sum_i g_i e^{-\beta\epsilon_i} \quad (3.16)$$

The exponential function $e^{-\beta\epsilon_i}$ is referred to as a *Boltzmann factor*. The probability of any quantum state being occupied is thus directly proportional to the Boltzmann factor for that state, which is a function of the energy and degeneracy of that state.

Fermi-Dirac Statistics

Maximizing Equation 3.9 yields

$$\frac{\partial}{\partial n_i} \left[\sum_i g_i \ln g_i - n_i \ln n_i - (g_i - n_i) \ln (g_i - n_i) \right] -$$

$$\alpha \frac{\partial}{\partial n_i} \left[\sum_{i=1}^N n_i - N \right] - \beta \frac{\partial}{\partial n_i} \left[\sum_{i=1}^N n_i \epsilon_i - \overline{E} \right] = 0$$

$$-\ln n_i - \frac{n_i}{n_i} + \ln (g_i - n_i) - \frac{g_i - n_i}{g_i - n_i} (-1) - \alpha - \beta\epsilon_i = 0$$

³The letter Z stands for *Zustandssumme*, which is German for “sum of states”.

$$\ln \left(\frac{g_i}{n_i} - 1 \right) = \alpha + \beta \epsilon_i$$

$$\frac{n_i}{g_i} = \frac{1}{e^{(\alpha + \beta \epsilon_i)} + 1} \quad (3.17)$$

Bose-Einstein Statistics

Maximizing Equation 3.10 yields

$$\begin{aligned} \frac{\partial}{\partial n_i} \left[\sum_i (n_i + g_i - 1) \ln (n_i + g_i - 1) - n_i \ln n_i \right] - \\ \frac{\partial}{\partial n_i} \left[\sum_i (g_i - 1) \ln (g_i - 1) \right] - \\ \alpha \frac{\partial}{\partial n_i} \left[\sum_{i=1}^N n_i - N \right] - \beta \frac{\partial}{\partial n_i} \left[\sum_{i=1}^N n_i \epsilon_i - \bar{E} \right] = 0 \\ \ln \left(\frac{n_i + g_i - 1}{n_i} \right) = \alpha + \beta \epsilon_i \end{aligned}$$

If we assume⁴ that $n_i + g_i \gg 1$, we have

$$\frac{n_i}{g_i} = \frac{1}{e^{(\alpha + \beta \epsilon_i)} - 1} \quad (3.18)$$

The distributions shown in Equation 3.13, Equation 3.17, and Equation 3.18 are referred to as the *canonical distribution* for the respective statistical method. Notice that since Equation 3.8 and Equation 3.11 differ by only the constant $N!$, the canonical distribution for Boltzmann statistics is identical to the canonical distribution for Maxwell-Boltzmann statistics. In other words, Equation 3.14 is valid for both Boltzmann statistics and Maxwell-Boltzmann statistics.

3.4 Alternative Derivation Of The Canonical Distribution

The probability of finding System A in a particular macrostate n with energy $E_{A,n}$ is proportional to the total multiplicity of the system when System A is in that macrostate. This probability is therefore proportional to the multiplicity of System B when the energy of System B is $E_B = E_{total} - E_{A,n}$.

$$p_{A,n} = \frac{\Omega_B(E_{total} - E_{A,n})}{\sum_n \Omega_B(E_{total} - E_{A,n})}$$

⁴This is justified by the fact there is no limit on the number of particles that can be found in any energy state. Thus, there are no limits on how large n_i can be.

3.4. ALTERNATIVE DERIVATION OF THE CANONICAL DISTRIBUTION 53

In writing this equation we are assuming that multiplicity is a function of the energy of the system. I'm not convinced there is any good way to justify this right now, so we're just going to go with it. We can use Equation 3.7 to express this equation in terms of the entropy of System B.

$$\mathcal{S} = \ln \Omega \quad \rightarrow \quad \Omega = e^{\mathcal{S}} \quad \rightarrow \quad p_{A,n} = \frac{e^{\mathcal{S}_B(E_{total} - E_{A,n})}}{\sum_n e^{\mathcal{S}_B(E_{total} - E_{A,n})}}$$

Next, we can express the entropy in this equation using a series expansion.

$$\mathcal{S}_B(E_{total} - E_{A,n}) = \mathcal{S}_B(E_{total}) + \left(\frac{\partial \mathcal{S}_B}{\partial E} \right) \bigg|_{E=E_{total}} (-E_{A,n}) + \dots$$

We then define the partial derivative at constant volume of entropy with respect to energy as the variable β .

$$\frac{\partial \mathcal{S}}{\partial E} = \beta \quad (3.19)$$

Thus,

$$\mathcal{S}_B(E_{total} - E_{A,n}) = \mathcal{S}_B(E_{total}) + \beta_B(-E_{A,n}) + \dots$$

Now let's assume that the energy of System B is much larger than the energy of System A. Under these circumstances, we refer to system B as an *energy reservoir* or *heat reservoir* for System A. The energy of the reservoir doesn't change when it exchanges energy with System A. Therefore, the energy of System B is approximately the same as the total energy of the system.

$$E_B \gg E_A \quad \rightarrow \quad E_B \approx E_{total}$$

Since entropy is a function of energy, we would also expect that

$$\mathcal{S}_B(E_{total} - E_{A,n}) \approx \mathcal{S}_B(E_{total})$$

and thus that we can truncate the series expansion introduced above for \mathcal{S}_B after the second term without introducing too much error⁵.

$$\mathcal{S}_B(E_{total} - E_{A,n}) = \mathcal{S}_B(E_{total}) - \beta_B E_{A,n}$$

We can now express the probability of finding System A in a particular macrostate with energy $E_{A,n}$ as

$$p_{A,n} = \frac{e^{\mathcal{S}_B(E_{total}) - \beta_B E_{A,n}}}{\sum_r e^{\mathcal{S}_B(E_{total}) - \beta_B E_{A,n}}} \quad \rightarrow \quad p_{A,n} = \frac{e^{\mathcal{S}_B(E_{total})} e^{-\beta_B E_{A,n}}}{e^{\mathcal{S}_B(E_{total})} \sum_r e^{-\beta_B E_{A,n}}}$$

⁵Of course, it's a separate conversation about how much is too much.

$$p_{A,n} = \frac{\cancel{e^{S_B(E_{total})}} e^{-\beta_B E_{A,n}}}{\cancel{e^{S_B(E_{total})}} \sum_r e^{-\beta_B E_{A,n}}} \rightarrow p_{A,n} = \frac{e^{-\beta_B E_{A,n}}}{\sum_r e^{-\beta_B E_{A,n}}}$$

$$p_{A,n} = \frac{e^{-\beta_B E_{A,n}}}{Z}$$

As defined previously, the exponential function $e^{-\beta E_{A,n}}$ is the Boltzmann factor for the state and the sum over states in the denominator is the partition function. We can therefore write the canonical distribution for systems obeying Boltzmann statistics as

$$p_n = \frac{e^{-\beta E_n}}{\sum_n e^{-\beta E_n}} \rightarrow p_n = \frac{e^{-\beta E_n}}{Z} \quad (3.20)$$

Let's pause for a moment to discuss this further. The partition function is a sum over the states accessible to the system. There are at least two ways we can compute that summation. We could sum over the states themselves (Equation 3.20) or we could sum over the energies of those states (Equation 3.16), in which case we must include the degeneracy of the energy states. For example, let's assume that we have a system with the following accessible energy states: ϵ , 2ϵ , 2ϵ , 2ϵ , 3ϵ , 3ϵ . When summing over the states accessible to the system the partition function is

$$Z = e^{-\beta\epsilon} + e^{-2\beta\epsilon} + e^{-2\beta\epsilon} + e^{-2\beta\epsilon} + e^{-3\beta\epsilon} + e^{-3\beta\epsilon}$$

$$Z = e^{-\beta\epsilon} + 3e^{-2\beta\epsilon} + 2e^{-3\beta\epsilon}$$

When summing over the energies of the accessible states we include the degeneracy of those states: $g_\epsilon = 1$, $g_{2\epsilon} = 3$, and $g_{3\epsilon} = 2$

$$Z = e^{-\beta\epsilon} + 3e^{-2\beta\epsilon} + 2e^{-3\beta\epsilon}$$

As expected, these approaches yield the same expression. Similarly, when calculating a probability for a state to be occupied using a Boltzmann factor, we include the degeneracy when expressing the probability in terms of the energy of the states (Equation 3.15) but do not include the degeneracy when expressing this probability in terms of the states themselves (Equation 3.20). For the accessible states in this example, the probability of being in a state with energy 2ϵ is

$$p_i = \frac{g_i e^{-\beta\epsilon_i}}{Z} \rightarrow p_{2\epsilon} = \frac{3e^{-2\beta\epsilon}}{Z}$$

But when expressing the probability with respect to the states themselves (Equation 3.20) we have

$$p_{2\epsilon} = \frac{e^{-2\beta\epsilon} + e^{-2\beta\epsilon} + e^{-2\beta\epsilon}}{Z} \rightarrow p_{2\epsilon} = \frac{3e^{-2\beta\epsilon}}{Z}$$

As expected, these approaches yield the same expression.

The equation for the probability of finding a system in the n^{th} macrostate (Equation 3.20) can be rearranged to form an equation for the energy E_n of that macrostate.

$$p_n = \frac{1}{Z} e^{-\beta E_n} \rightarrow E_n = -\frac{1}{\beta} (\ln(Z) + \ln(p_n)) \quad (3.21)$$

The probability distribution in Equation 3.21 is also generally referred to as the canonical distribution, although through comparison to Equation 3.14 we see that it is specifically the canonical distribution for systems obeying Boltzmann statistics or Maxwell-Boltzmann statistics.

We can use Equation 3.21 to rewrite the expression for dQ in Equation 2.6.

$$\begin{aligned} \sum_n (dp_n) E_n &= \sum_n (dp_n) \left(-\frac{1}{\beta} \right) (\ln(Z) + \ln(p_n)) \\ \sum_n (dp_n) E_n &= -\frac{1}{\beta} \left[\sum_n \ln(Z) (dp_n) + \sum_n (dp_n) \ln(p_n) \right] \end{aligned}$$

Since $\ln(Z)$ is a constant, it can be taken outside of the summation.

$$\sum_n (dp_n) E_n = -\frac{1}{\beta} \left[\ln(Z) \sum_n dp_n + \sum_n (dp_n) \ln(p_n) \right]$$

We now recognize that if the distribution of probabilities p_n is normalized, there is a constraint on the sum of dp_n ⁶.

$$\sum_n p_n = 1 \rightarrow \sum_n dp_n = 0 \rightarrow \sum_n (dp_n) E_n = -\frac{1}{\beta} \sum_n (dp_n) \ln(p_n)$$

Thus,

$$dQ = -\frac{1}{\beta} \sum_n (dp_n) \ln(p_n) \quad (3.22)$$

Let's pause here for a second and use Equation 3.19 to rewrite this expression

$$\begin{aligned} \beta dQ &= -\sum_n (dp_n) \ln(p_n) \\ \left(\frac{\partial S}{\partial E} \right) dQ &= -\sum_n (dp_n) \ln(p_n) \end{aligned}$$

⁶Needless to say, we will always assume that the probability distribution is normalized.

The left hand side of this expression thus denotes a change in entropy associated with the energy change of the system described by the absorption (or release) of heat. In light of this, the right hand side of this expression must thus represent a change in entropy of the system, as we will discuss later on.

3.5 Entropy and Heat

Let's assume that multiplicity is a function of energy and volume. If so, then entropy must also be a function of energy and volume.

$$S = S(E, V) \rightarrow dS = \left(\frac{\partial S}{\partial E} \right)_V dE + \left(\frac{\partial S}{\partial V} \right)_E dV$$

As mentioned before, the subscripts indicate which variables are held constant in the partial derivative. For example, $\left(\frac{\partial S}{\partial E} \right)_V$ is the partial derivative of the entropy with respect to energy when the volume is held constant. Substitution of the previous definition of β gives us

$$dS = \beta dE + \left(\frac{\partial S}{\partial V} \right)_E dV \rightarrow dE = \frac{1}{\beta} dS - \frac{1}{\beta} \left(\frac{\partial S}{\partial V} \right)_E dV$$

The second term on the right is a change in energy associated with a change in the volume of the system. This can therefore be equated with work.

$$\frac{1}{\beta} \left(\frac{\partial S}{\partial V} \right)_E dV = dW \rightarrow dE = \frac{1}{\beta} dS - dW$$

We thus conclude from Equation 1.2 that

$$\frac{1}{\beta} dS = dQ \rightarrow dS = \beta dQ \quad (3.23)$$

Equation 3.23 is one way to express the 2nd law of thermodynamics. Let's now combine this result with Equation 3.22 to derive an expression for dS .

$$dQ = -\frac{1}{\beta} \sum_n (dp_n) \ln(p_n) \rightarrow dS = -\sum_n (dp_n) \ln(p_n)$$

We now embark on some algebra, starting with subtracting $0 = \sum_n dp_n$ from both sides of the equation⁷.

$$dS - 0 = -\sum_n (dp_n) \ln(p_n) - \sum_n dp_n$$

We now work to write the right side of this equation as a total derivative.

$$dS = -\sum_n (dp_n) \ln(p_n) - \sum_n p_n \left(\frac{dp_n}{p_n} \right)$$

⁷We're again assuming that the probability distribution is normalized.

$$\begin{aligned}
d\mathcal{S} &= - \sum_n \left[(dp_n) \ln(p_n) + p_n \left(\frac{dp_n}{p_n} \right) \right] \\
d\mathcal{S} &= - \sum_n [(dp_n) \ln(p_n) + p_n d(\ln(p_n))] \\
d\mathcal{S} &= d \left[- \sum_n p_n \ln(p_n) \right]
\end{aligned}$$

Thus,

$$\mathcal{S} = - \sum_n p_n \ln(p_n) \quad (3.24)$$

This definition of entropy comes from the role played by heat in energy conservation and then back to Equation 3.22 that relates heat to probability. However this equation is predicated on the canonical distribution, which assumes that our system is in contact with an energy reservoir that maintains a constant value of β ; we also assume that this system can be described using Boltzmann or Maxwell-Boltzmann statistics. So, it would therefore seem that Equation 3.24 is valid only when these conditions are met and thus that Equation 3.24 is less general than Equation 3.7.

We recognize from Equation 2.7 and Equation 3.24 that

$$\mathcal{S} = -H$$

Thus, from Boltzmann's H -theorem (Equation 2.11) we conclude that

$$\frac{d\mathcal{S}}{dt} \geq 0 \quad (3.25)$$

The entropy of a closed system governed by Hamiltonian mechanics must increase with time. Only when the system is in equilibrium will the entropy be constant in time. This is yet another way of expressing the 2nd law of thermodynamics. It is arguably more common to express the 2nd law of thermodynamics as $\frac{d\mathcal{S}}{dt} \geq 0$ than as $\frac{1}{\beta} d\mathcal{S} = dQ$.

Conjugate Pairs Again

We see from Equation 3.23 that $\frac{1}{\beta}$ and \mathcal{S} are also a conjugate pair (Section 1.5). In this case, $\frac{1}{\beta}$ is the intensive variable (*i.e.*, the generalized force) and \mathcal{S} is the extensive variable (*i.e.*, the generalized displacement). This allows us to write the 1st law of thermodynamics as

$$dE = \frac{d\mathcal{S}}{\beta} - PdV + \mu dN$$

The parameter $\frac{1}{\beta}$ thus reflects the change in the energy of the system associated with a change in the entropy of the system. Furthermore, we see that

$$\frac{1}{\beta} = \left(\frac{\partial E}{\partial S} \right)_{V,N} \quad P = \left(\frac{\partial E}{\partial V} \right)_{S,N} \quad \mu = \left(\frac{\partial E}{\partial N} \right)_{S,V}$$

This form of the 1st law of thermodynamics can also be written as

$$dS = \beta dE + \beta P dV - \beta \mu dN$$

Thus,

$$\beta = \left(\frac{\partial S}{\partial E} \right)_{V,N} \quad \beta P = \left(\frac{\partial S}{\partial V} \right)_{E,N} \quad \beta \mu = - \left(\frac{\partial S}{\partial N} \right)_{E,V} \quad (3.26)$$

One final note about this. Since energy and entropy are exact differentials, their second mixed partial derivatives must be equal. For example,

$$\frac{\partial^2 E}{\partial V \partial N} = \frac{\partial^2 E}{\partial N \partial V} \quad \rightarrow \quad \left(\frac{\partial}{\partial V} \right) \left(\frac{\partial E}{\partial N} \right)_{S,V} = \left(\frac{\partial}{\partial N} \right) \left(\frac{\partial E}{\partial V} \right)_{S,N}$$

Substitution of the previously derived definitions then yields

$$\left(\frac{\partial}{\partial V} \right) (\mu) = \left(\frac{\partial}{\partial N} \right) (P) \quad \rightarrow \quad \left(\frac{\partial \mu}{\partial V} \right)_N = \left(\frac{\partial P}{\partial N} \right)_V$$

This equation is an example of a collection of identities referred to collectively as Maxwell Relations. We will discuss this further in Section 6.5.5

3.6 Boltzmann's Model For An Ideal Gas

In 1877 our friend Boltzmann published a paper⁸ on ideal gases that used statistical approaches to describe their behavior and properties. He began assuming that each molecule of the gas can have only a finite number of possible values for its energy.

$$0, \epsilon, 2\epsilon, 3\epsilon \dots q\epsilon$$

The constant q in this equation is an arbitrary finite number. It is worth pausing at this point to acknowledge two things. First, Boltzmann had to assume that the accessible energy states for the ideal gas are effectively quantized because, as we shall see, he needs to compute a summation over them. Second, we should remember that this statement (or assumption) predates any discussion of quantum mechanics by several decades. It goes without saying that this assumption caused some controversy since kinetic energy was believed to be continuous just as speed was continuous. Boltzmann skirted this controversy

⁸On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of Heat and Probability Calculations Regarding the Conditions for Thermal Equilibrium

by allowing ϵ to infinitesimally small, which then effectively recapitulates a continuous distribution of possible kinetic energies for the molecules in the gas.

Following Boltzmann's derivation, we further assume that both the number of molecules of the gas, denoted by N , and the total energy of the gas, denoted by E , are both constant⁹.

$$N = n_0 + n_1 + n_2 + \cdots n_q$$

$$E = n_0(0) + n_1(\epsilon) + n_2(2\epsilon) + \cdots n_q(q\epsilon)$$

The variable n_i in these equations denotes the number of molecules of the gas with an energy $i\epsilon$. The permutation of distributing the N molecules of gas into these different energy states is

$$P = \frac{N!}{n_0!n_1!n_2!\cdots n_q!}$$

$$\ln(P) = \ln(N!) - \ln(n_0!) - \ln(n_1!) - \cdots - \ln(n_q!)$$

We can simplify this equation using Stirling's approximation.

$$\ln P \approx N \ln(N) - N - [n_0 \ln(n_0) - n_0 + n_1 \ln(n_1) - n_1 + \cdots + n_q \ln(n_q) - n_q]$$

$$\ln P \approx N \ln(N) - [n_0 \ln(n_0) + n_1 \ln(n_1) + \cdots + n_q \ln(n_q)] + [n_0 + n_1 + \cdots n_q - N]$$

The final term in this expression is zero.

$$\ln P \approx N \ln(N) - [n_0 \ln(n_0) + n_1 \ln(n_1) + \cdots + n_q \ln(n_q)]$$

$$\ln P \approx N \ln(N) - \sum_{j=0}^q n_j \ln(n_j)$$

The distribution of molecules (*i.e.*, distribution of n_0, n_1, \dots, n_q) observed at equilibrium is the distribution corresponding to the largest value of P ; the justification for this statement is similar to what we used in Section 3.1 to argue that systems are most likely to be found in their most probable configuration. Thus, the distribution of molecules observed at equilibrium is the distribution corresponding to when $-\sum_{j=0}^q n_j \ln(n_j)$ is a maximum.

Let's now assume that ϵ is very small; this also helps us address the "energy should be continuous" complaining. In this case, we can then regard all particles with energy values between x and $x + \epsilon$ as having the same kinetic energy. We

⁹We can assure that the energy of the system is constant by making the system isolated from the rest of the universe.

thus define the variable n_0 to be the number of particles with an energy between 0 and ϵ , the variable n_1 to be the number of particles with an energy between ϵ and 2ϵ , *etc.* Since the variables $n_0, n_1, \text{etc.}$ are all therefore infinitesimally small and on the order of the magnitude of ϵ , we can write them in terms of a distribution, $f(x)$, that describes the probability of a particle having energy x .

$$n_0 = \epsilon f(0) \quad n_1 = \epsilon f(\epsilon) \quad n_2 = \epsilon f(2\epsilon) \quad \text{etc}$$

We can thus express the sum to be maximized in terms of the distribution $f(\epsilon)$ ¹⁰.

$$\begin{aligned} -\sum_{j=0}^q n_j \ln(n_j) &= -\sum_{j=0}^q \epsilon f(j\epsilon) \ln[\epsilon f(j\epsilon)] \\ -\sum_{j=0}^q n_j \ln(n_j) &= -\sum_{j=0}^q \epsilon f(j\epsilon) (\ln[\epsilon] + \ln[f(j\epsilon)]) \\ -\sum_{j=0}^q n_j \ln(n_j) &= -\sum_{j=0}^q \epsilon f(j\epsilon) (\ln[\epsilon]) - \sum_{j=0}^q \epsilon f(j\epsilon) (\ln[f(j\epsilon)]) \\ -\sum_{j=0}^q n_j \ln(n_j) &= -(\ln[\epsilon]) \sum_{j=0}^q \epsilon f(j\epsilon) - \sum_{j=0}^q \epsilon f(j\epsilon) (\ln[f(j\epsilon)]) \end{aligned}$$

The first summation on the right-hand side of this equation is thus the sum over the number of particles in the energy states.

$$\begin{aligned} -\sum_{j=0}^q n_j \ln(n_j) &= -(\ln[\epsilon]) \sum_{j=0}^q n_j - \sum_{j=0}^q \epsilon f(j\epsilon) (\ln[f(j\epsilon)]) \\ -\sum_{j=0}^q n_j \ln(n_j) &= -(\ln[\epsilon]) N - \sum_{j=0}^q \epsilon f(j\epsilon) (\ln[f(j\epsilon)]) \end{aligned}$$

The first term on the right-hand side of this equation is a constant and therefore irrelevant to our maximization problem. We are thus left with maximizing the second term.

$$-\sum_{j=0}^q \epsilon f(j\epsilon) (\ln[f(j\epsilon)])$$

As we make ϵ smaller and smaller, we approach a continuum allowing us to replace this summation by an integration. We denote this quantity by the variable \mathcal{S} as it is equivalent to the definition of \mathcal{S} in Equation 3.24.

$$\mathcal{S} = - \int f(\epsilon) \ln[f(\epsilon)] d\epsilon \quad (3.27)$$

¹⁰Incidentally, this derivation indicates that the units of $f(\epsilon)$ are number of particles per energy.

Boltzmann referred to this quantity as the permutability measure¹¹, but we now refer to it as entropy. Indeed, it is the integral form of the summation in Equation 3.24.

Wait a minute

We arrived at Equation 3.27 from the multiplicity of the molecules of the gas, and thus eventually from Equation 3.7, under the assumption that the number of molecules in the gas is so large that the Stirling's approximation was valid. Alternatively, we arrive at Equation 3.24 from the canonical distribution, which assumed that the system was in contact with an energy reservoir at constant β . And yet, these two expressions (Equation 3.24 and Equation 3.27) for entropy are (effectively) the same.

The Canonical Distribution, Again

We can now explore what function $f(\epsilon)$ would maximize the entropy. More specifically, we seek a function $f(\epsilon)$ that maximize Equation 3.27 subject to the following two constraints:

$$1 = \int_0^{\infty} f(\epsilon) d\epsilon \quad \bar{\epsilon} = \int_0^{\infty} \epsilon f(\epsilon) d\epsilon$$

We recognize these constraints as being the zeroth moment and first moment of the distribution $f(\epsilon)$ ¹². For this particular distribution of the molecules, the zeroth moment is constrained by the probability function being normalized and the first moment is constrained by the mean (kinetic) energy per molecule of the gas, denoted as $\bar{\epsilon}$. We can determine the value of $f(\epsilon)$ subject to these two constraints using our old friend Lagrange multipliers.

$$-\int_0^{\infty} f(\epsilon) \ln[f(\epsilon)] d\epsilon - \alpha \left(\int_0^{\infty} f(\epsilon) d\epsilon - 1 \right) - \beta \left(\int_0^{\infty} \epsilon f(\epsilon) d\epsilon - \bar{\epsilon} \right) = \text{maximum}$$

Differentiating with respect to $f(\epsilon)$ then gives

$$-1 - \ln(f(\epsilon)) - \alpha - \beta\epsilon = 0 \quad \rightarrow \quad f(\epsilon) = e^{-1-\alpha} e^{-\beta\epsilon}$$

$$f(\epsilon) = C e^{-\beta\epsilon} \tag{3.28}$$

We recognize this as the canonical distribution, yet again.

¹¹It sounds better in the original German - Permutabilitätssmass

¹²See Section 3.1.1

3.7 So What Is Entropy?

Entropy is a number that reports on the statistical nature of a system or distribution. Three common expressions for calculating it are shown in Equation 3.29.

$$\begin{aligned}
 \mathcal{S} &= \ln \Omega \\
 \mathcal{S} &= - \sum_n p_n \ln(p_n) = -\overline{\ln p} \\
 \mathcal{S} &= - \int p(x) \ln(p(x)) dx = -\overline{\ln p}
 \end{aligned} \tag{3.29}$$

Let's consider a one-dimensional box of length a . A uniform and normalized probability distribution within that box can be described by the following equation

$$p(x) = \begin{cases} \frac{1}{a} & 0 < x < a \\ 0 & elsewhere \end{cases}$$

We can calculate the entropy of this distribution using the integral expression in Equation 3.29.

$$\begin{aligned}
 \mathcal{S} &= - \int_0^a \frac{1}{a} \ln\left(\frac{1}{a}\right) dx \\
 \mathcal{S} &= -\ln\left(\frac{1}{a}\right) \rightarrow \mathcal{S} = \ln(a)
 \end{aligned}$$

The equivalent multiplicity of this system can be found using Equation 3.7.

$$\Omega = e^{\ln(a)} \rightarrow \Omega = a$$

The multiplicity reflects the scale of the non-zero portion of the probability distribution. The larger this size, the larger the multiplicity.

Let's now consider a probability distribution that is narrower and higher, but still normalized.

$$p(x) = \begin{cases} \frac{2}{a} & \frac{a}{2} < x < a \\ 0 & elsewhere \end{cases}$$

The entropy associated with this distribution is

$$\mathcal{S} = - \int_{\frac{a}{2}}^a \frac{2}{a} \ln\left(\frac{2}{a}\right) dx$$

$$\mathcal{S} = -\ln\left(\frac{2}{a}\right) \rightarrow \mathcal{S} = \ln\left(\frac{a}{2}\right)$$

Note that since $\ln\left(\frac{a}{2}\right) < \ln(a)$, the entropy of the narrower and higher distribution is less than the entropy of the wider and lower distribution. We could argue that this corresponds to the former distribution being more localized. In other words, entropy is a measure of the localization of a probability distribution. The more localized the distribution, the lower the entropy.

3.8 Independence

The definition of entropy in Equation 3.29 automatically gives rise to a few interesting properties for entropy as a quantity. First, let's consider two systems, System A and System B. System A has a probability p_r of being found in state r and System B has a probability p_s of being found in state s . The entropies of these two systems are

$$\mathcal{S}_A = -\sum_r p_r \ln(p_r) \quad \mathcal{S}_B = -\sum_s p_s \ln(p_s)$$

Similarly, we can define the entropy of the composite system consisting of System A plus System B as

$$\mathcal{S} = -\sum_s \sum_r p_{rs} \ln(p_{rs})$$

In this equation, p_{rs} denotes the probability of finding System A in state r and System B in state s simultaneously. This composite probability p_{rs} is related to the individual probabilities for each system p_r and p_s by the following equations

$$p_s = \sum_r p_{rs} \quad p_r = \sum_s p_{rs}$$

Let's now calculate the difference between the entropy of the composite system and the sum of the entropies of the individual systems.

$$\mathcal{S} - (\mathcal{S}_1 + \mathcal{S}_2) = -\sum_s \sum_r p_{rs} \ln(p_{rs}) - \left(-\sum_r p_r \ln(p_r) - \sum_s p_s \ln(p_s) \right)$$

We can rewrite the final two summations in this expression in terms of p_{rs} so that we have a common term for all summations.

$$\mathcal{S} - (\mathcal{S}_1 + \mathcal{S}_2) = -\sum_s \sum_r p_{rs} \ln(p_{rs}) - \left(-\sum_r \sum_s p_{rs} \ln(p_r) - \sum_s \sum_r p_{rs} \ln(p_s) \right)$$

And now we simplify

$$\mathcal{S} - (\mathcal{S}_1 + \mathcal{S}_2) = - \sum_s \sum_r [p_{rs} \ln(p_{rs}) - p_{rs} \ln(p_r) - p_{rs} \ln(p_s)]$$

$$\mathcal{S} - (\mathcal{S}_1 + \mathcal{S}_2) = - \sum_s \sum_r \left[p_{rs} \ln \left(\frac{p_{rs}}{p_r p_s} \right) \right]$$

Let's switch the numerator and the denominator in the natural log term to get rid of that negative sign.

$$\mathcal{S} - (\mathcal{S}_1 + \mathcal{S}_2) = \sum_s \sum_r \left[p_{rs} \ln \left(\frac{p_r p_s}{p_{rs}} \right) \right]$$

Finally, since $\ln(x) \leq x - 1$, we have

$$\ln \left(\frac{p_r p_s}{p_{rs}} \right) \leq \frac{p_r p_s}{p_{rs}} - 1$$

Thus,

$$\mathcal{S} - (\mathcal{S}_1 + \mathcal{S}_2) \leq \sum_s \sum_r \left[p_{rs} \left(\frac{p_r p_s}{p_{rs}} - 1 \right) \right]$$

$$\mathcal{S} - (\mathcal{S}_1 + \mathcal{S}_2) \leq \sum_s \sum_r [p_r p_s - p_{rs}]$$

The entropy of the combined system is therefore always less than or equal to the sum of the two separate systems. The equality holds if the systems are statistically independent, in which case $p_{rs} = p_r p_s$ and

$$\mathcal{S} - (\mathcal{S}_1 + \mathcal{S}_2) = 0 \quad \rightarrow \quad \mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$$

Thus, the entropy of a system consisting of statistically independent systems is the sum of the entropies of those systems.

3.9 Summary

- The multiplicity (Ω) of a macrostate is the number of microstates associated with the macrostate.
- An isolated system in equilibrium is equally likely to be in any of its accessible states, each of which is defined by a particular configuration of the system's elements.
- The probability of finding an isolated system in equilibrium in a particular macrostate is directly proportional to the multiplicity of that macrostate. In other words, it is directly proportional to the number of microstates associated with the macrostate.

- Entropy is a number that reports on the statistical nature of a system or distribution.

$$\begin{aligned}\mathcal{S} &= \ln \Omega \\ \mathcal{S} &= - \sum_n p_n \ln(p_n) = -\overline{\ln p} \\ \mathcal{S} &= - \int p(x) \ln(p(x)) dx = -\overline{\ln p}\end{aligned}$$

- The pressure and chemical potential of system can be determined from the equation for the entropy for that system.

$$\beta = \left(\frac{\partial \mathcal{S}}{\partial E} \right)_{V,N} \quad \beta P = \left(\frac{\partial \mathcal{S}}{\partial V} \right)_{E,N} \quad \beta \mu = - \left(\frac{\partial \mathcal{S}}{\partial N} \right)_{V,N}$$

- The Second Law of Thermodynamics states that the entropy of any isolated system will increase with time.
- Heat is associated with changes to the probability that a state is occupied

$$dQ = -\frac{1}{\beta} \sum_n (dp_n) \ln(p_n)$$

- The **partition function** is the summation over the states accessible to a system. It can be written as a summation over the states or as a summation over the energies of those states.

$$Z = \sum_{\substack{n \\ \text{states}}} e^{-\beta \epsilon_n} \quad Z = \sum_{\substack{i \\ \text{energies}}} g_i e^{-\beta \epsilon_i}$$

- The **canonical distribution** is the probability distribution associated with a system where the number of particles and mean energy are fixed or otherwise constrained.

$$f(\epsilon) = C e^{-\beta \epsilon}$$

Boltzmann distribution

Classical particles that are distinguishable.

- $\Omega_B = N! \prod_i \frac{g_i^{n_i}}{n_i!}$
- $\mathcal{S}_B = \left(\frac{N}{e} \right)^N + \sum_i n_i + \sum_i n_i \ln \left(\frac{g_i}{n_i} \right)$
- Maximum entropy distribution: $\frac{n_i}{g_i} = \frac{N e^{-\beta \epsilon_i}}{\sum_i g_i e^{-\beta \epsilon_i}}$

Maxwell-Boltzmann distribution

Classical particles that are indistinguishable.

- $\Omega_{MB} = \prod_i \frac{g_i^{n_i}}{n_i!}$
- $\mathcal{S}_{MB} = \sum_i n_i - \sum_i n_i \ln \left(\frac{n_i}{g_i} \right)$
- Maximum entropy distribution: $\frac{n_i}{g_i} = \frac{N e^{-\beta \epsilon_i}}{\sum_i g_i e^{-\beta \epsilon_i}}$

Fermi-Dirac distribution

Quantum mechanical particles that are identical, indistinguishable, and subject to the Pauli Exclusion Principle.

- $\Omega_{FD} = \prod_i \frac{g_i!}{n_i!(g_i - n_i)!}$
- $\mathcal{S}_{FD} = \sum_i g_i \ln g_i - n_i \ln n_i - (g_i - n_i) \ln (g_i - n_i)$
- Maximum entropy distribution: $\frac{n_i}{g_i} = \frac{1}{e^{(\alpha + \beta \epsilon_i)} + 1}$

Bose-Einstein distribution

Quantum mechanical particles that are identical, indistinguishable, and not subject to the Pauli Exclusion Principle.

- $\Omega_{BE} = \prod_i \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}$
- $\mathcal{S}_{BE} = \sum_i (n_i + g_i - 1) \ln (n_i + g_i - 1) - n_i \ln n_i - (g_i - 1) \ln (g_i - 1)$
- Maximum entropy distribution: $\frac{n_i}{g_i} = \frac{1}{e^{(\alpha + \beta \epsilon_i)} - 1}$