

The second law of thermodynamics - part I.

Asaf Pe'er¹

September 20, 2013

1. The direction of natural processes: basic ideas

We know from everyday life that when left to itself, a system which is initially not in equilibrium always changes towards equilibrium, while the opposite direction does not happen. There is nothing in the first law that prevents the system to move in the opposite direction, since energy is equally conserved.

Examples.

1. Take two bodies, one hotter than the other and put them in contact, - *heat will flow from the hot body to the cold one, but not the opposite direction.*

2. Consider a gas that fills a container. About half of the molecules (but never exactly half !) are in one half of the container, the rest in the second half. Each molecule can move freely, so in principle nothing prevents all the molecules to occupy only 1/2 of the container. But left alone, it never happens (never = the time it takes this to happen is longer than the age of the universe). But the opposite direction does happen: if we begin by filling only 1/2 of the container with molecules, after some time they fill the entire container. This process is **irreversible**. **We lost information during this process:** initially, we knew that all the molecules are in one half, but at the end we cannot constrain the location of each molecule inside the container. The final state is thus *less ordered* or *more random* than the initial one.

The **entropy** of the state a system is **a quantitative measure of the degree of disorder of that state**. More disorder = larger entropy. *The entropy is a function of state.* Entropy provides a quantitative criterion for the direction of processes in nature.

2. Statistical weight

Consider a system consisting of N molecules from only one type. We discriminate between the macroscopic state of the system (in short: **macrostate**), and the microscopic state of the system (**microstate**).

¹Physics Dep., University College Cork

A **macrostate** is determined by the values of macroscopic parameters, such as the volume, V , energy E or number of particles, N . A **microstate** is determined by the quantities (position, velocity, etc.) of all individual molecules (of the order of 10^{23} ...).

The microstate thus obviously contains much more information than the macrostate. However, practically, it contains too much information than can be handled. There are numerous number of microstates corresponding to a single macrostate. For example, a single macrostate corresponds to energy E , but this can be achieved when the molecules are in different locations (\vec{r}_i).

We can, though, think of counting microstates: maybe easiest to think of it in terms of quantum mechanics, where the microstates are a discrete set. Thus, *every macrostate of a system comprises a well defined number of microstates of that system*. We call this number the **statistical weight** of the system, and denote it by $\Omega = \Omega(E, V, N)$.

2.1. Example: paramagnetic solid in a magnetic field.

Let us take a simple example: a **paramagnetic** solid is composed of N molecules, each have a magnetic dipole moment $\vec{\mu}$ (these can be thought of as tiny magnets). In the absence of magnetic field, these dipoles are randomly oriented. However, when a magnetic field is applied, the dipoles align themselves either parallel or anti-parallel to the field. (We limit the discussion here only to the interaction between the dipole and the magnetic field, neglecting all other things such as vibrational motion of the molecules, etc. These will be discussed later).

The energy of a dipole in a magnetic field is

$$E = -\vec{\mu} \cdot \vec{B}. \quad (1)$$

Thus, if the dipole is oriented parallel to the magnetic field, its energy is $E = -\mu B$, while if it is anti-parallel, its energy is $E = +\mu B$.

If n dipoles are oriented parallel to the magnetic field, while the remaining $N - n$ are anti-parallel, the total energy of the system is

$$E(n) = -\mu B n + \mu B (N - n) = (N - 2n)\mu B. \quad (2)$$

Thus, n determines the energy of the system - its **macrostate**: it tells us *how many* dipoles are aligned with the field. However, n does *not* determines the **microstate**, which tells us *which* of the dipoles is aligned with the field. This is demonstrated by a simple system with $N = 9$ dipoles presented in Figure 1.

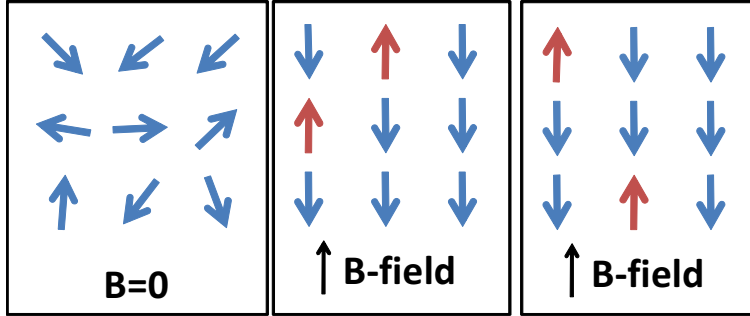


Fig. 1.— Array of $N = 9$ spin $1/2$ dipoles. Left: when there is no magnetic field, the dipoles are randomly aligned. Middle and right: two *different* microscopic states, both have $n = 2$ dipoles aligned parallel to the magnetic field and $(N - n) = 7$ dipoles aligned anti-parallel to the magnetic field. Thus, both correspond to *the same* macroscopic state.

We have in the system N dipoles, each having 2 possible orientations. Thus, there are total 2^N microstates. The number of microstates with energy $E(n)$ is equal to the number of ways in which n sites (filled with parallel dipoles) can be selected out of the total N available sites. The other $N - n$ sites are necessarily filled with anti-parallel dipoles. This number is

$$\Omega(n) = \binom{N}{n} \equiv \frac{N!}{n!(N-n)!} \quad (3)$$

If $n = N$, all dipoles are aligned parallel to the magnetic field. The energy is at its minimum: $E = (N - 2N)\mu B = -\mu BN$; this state is known as the “ground state”. The statistical weight of this state is $\Omega(N) = 1$: there is only one such microstate.

However, if $n = N/2$, $E(n = N/2) = (N - 2(N/2))\mu B = 0$. The system behaves as if no field applies. The statistical weight attains its maximum:

$$\Omega\left(n = \frac{N}{2}\right) = \frac{N!}{\left(\frac{N}{2}\right)!\left(\frac{N}{2}\right)!} \quad (4)$$

This weight rapidly grows with large N (see Figure 2). As we see, the statistical weight Ω is a measure of the order or disorder of a system. The minimum value ($\Omega = 1$) corresponds to maximum order (system being in a unique microstate). In the macrostate with maximum Ω , we know very little about the microstate of the system: this is the macrostate with maximum disorder.

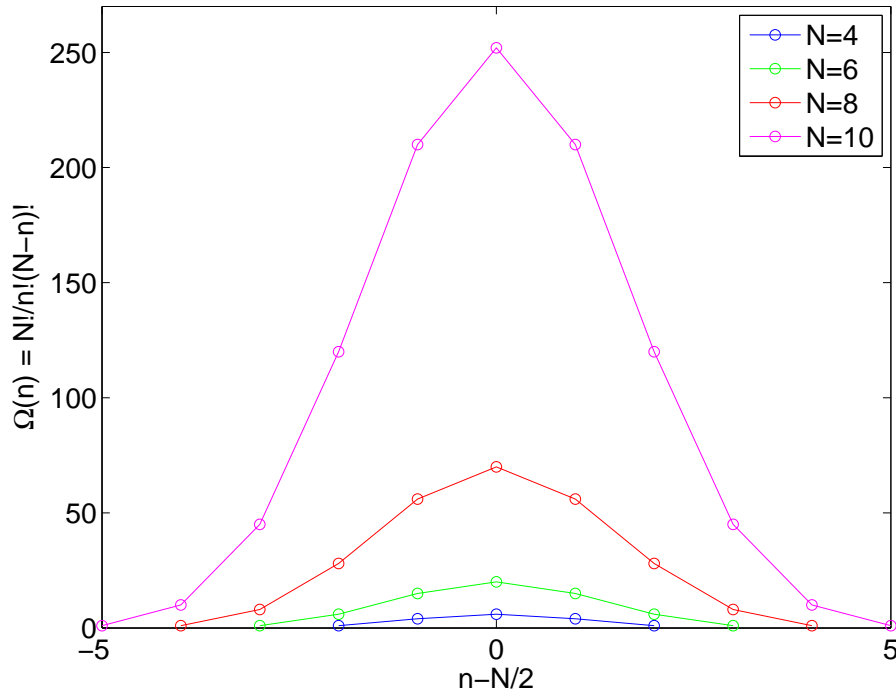


Fig. 2.— A plot of the statistical weight $\Omega(n)$ as a function of n , for several values of N . Clearly, while the maximum of $\Omega(n)$ is always obtained for $n = N/2$, where N grows, the ‘peak’ is much more pronounced. Many more microscopic states correspond to the same macroscopic state of $n = N/2$.

3. Equilibrium in an isolated system

Consider an isolated system. We would like to address the following question: *what are the conditions in which this system is in equilibrium?*

We know that in equilibrium, the state of the system is fully specified by the functions of state, (E, V, N) . Thus, if the system is *not* in equilibrium, we need to add (at least) one more variable, call it α . To each choice of α , there corresponds a statistical weight, $\Omega(E, V, N, \alpha)$ = the number of microstates corresponding to the given macrostate.

We now postulate the *equal a-priori probability postulate*, which states that **for an isolated system, all microstates compatible with the given constraints (E, V, N) are equally likely to occur.**

When the system is in equilibrium, α gets a particular value. We now make a second postulate, the *equilibrium postulate*: **The equilibrium corresponds to the value of α**

for which $\Omega(E, V, N, \alpha)$ attains its maximum value, for fixed (E, V, N) .

Physically, we claim that the equilibrium state is the state of maximum probability. Thus, while the system can be in many states, *it is much more probable* that it will be in the equilibrium state, simply because there are very many more microstates that correspond to this state. It is extremely rare that the system will be in a different state (unless disturbed from the outside).

We shall now replace the statistical weight Ω by a new quantity, the **entropy**, which will be *defined* by

$$\boxed{S \equiv k_B \ln \Omega} \quad (5)$$

Using Equation 5, we can now write the equilibrium postulate in terms of the entropy: **during a process the entropy of an isolated system increases (or remains, if the process is reversible). In a state of equilibrium, the entropy obtains its maximum value.** This statement is the basic form of **the second law of thermodynamics**.

We would now like to understand this seemingly strange form of the definition of the entropy: why did we use “ln” ? To see this, consider an isolated system, divided into two parts, labeled “1” and “2” (see Figure 3). Each part (subsystem) has its own energy, volume, and number of particles. Clearly, the total energy, volume and number of particles in the system is the sum of the two sub-systems:

$$\begin{aligned} E_T &= E_1 + E_2 \\ V_T &= V_1 + V_2 \\ N_T &= N_1 + N_2. \end{aligned} \quad (6)$$

In order for equilibrium to establish, the two sub-systems have to be able to interact; in particular, it is essential that energy exchange can occur. The statistical weight of the system is given by

$$\Omega_T(E_T, V_T, N_T) = \Omega_1(E_1, V_1, N_1) \times \Omega_2(E_2, V_2, N_2) \quad (7)$$

Equation 7 states the obvious fact that *every microstate of sub-system 1, can be combined with any microstate of sub-system 2*. Hence the multiplicity. Note that the conditions in Equation 6 imply that E_2, V_2 and N_2 are not independent.

We can now write the entropy of the total system by

$$S_T = k_B \ln \Omega_T = k_B \ln(\Omega_1 \Omega_2) = k_B \ln \Omega_1 + k_B \ln \Omega_2 = S_1 + S_2, \quad (8)$$

We therefore find that the entropy of the combined system is the sum of the entropies of the two sub-systems. Thus, **entropy is additive**. This is a direct consequence of the “ln” used in the definition of the entropy.

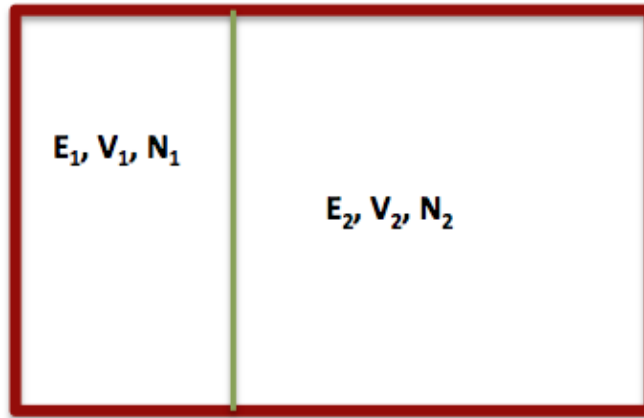


Fig. 3.— An isolated system divided into two parts

The two definitions of equilibrium state. We see that we can define an equilibrium state in two ways:

1. An equilibrium state is a state in which the system is in a macrostate with the highest statistical weight.
2. An equilibrium is a state when the entropy obtains its maximum value.

Equations 6, 8, imply that the quantities E , N , V and S are additive: they are proportional to the size of the system. If the system doubles its size - these quantities double. Quantities which are proportional to the size of the system are called **extensive** variables. As opposed to extensive variables, one can define **intensive** variables as variables which are independent on the size of the system. An example of intensive variables are temperature and pressure.

Example - 1.

Let us return to Figure 3, and consider the partition between the two parts of the system to be fixed and permeable to atoms, but allows heat to pass. Such a partition is called *diathermal wall*. Thus, V_1, V_2, N_1 and N_2 are fixed. There is only one independent variable - E_1 , since $E_2 = E_T - E_1$.

The equilibrium condition for the system is that the entropy is maximized. This condi-

tion is written as

$$\begin{aligned} \left(\frac{\partial S_T}{\partial E_1}\right)_{E_T=const} &= \left(\frac{\partial S_1}{\partial E_1}\right) + \left(\frac{\partial S_2}{\partial E_1}\right) \\ &= \left(\frac{\partial S_1}{\partial E_1}\right) + \left(\frac{\partial S_2}{\partial E_2}\right) \left(\frac{\partial E_2}{\partial E_1}\right) \\ &= \left(\frac{\partial S_1}{\partial E_1}\right) - \left(\frac{\partial S_2}{\partial E_2}\right) = 0. \end{aligned} \tag{9}$$

In the last equation we used the fact that $E_2 = E_T - E_1$, thus $\partial E_2/\partial E_1 = -1$. We thus find that the condition for equilibrium is obtained if

$$\left(\frac{\partial S_1}{\partial E_1}\right) = \left(\frac{\partial S_2}{\partial E_2}\right) \tag{10}$$

Equation 10 is the condition for thermal equilibrium. In this state, there is no heat transfer between the two sub-systems. In other words, the two sub-systems must have the *same temperature*. Thus, $\partial S_i/\partial E_i$ is a measure of the temperature of the sub-system i . We can use this insight to *define* an absolute temperature, T_i for each sub-system, by

$$\frac{1}{T_i} \equiv \left(\frac{\partial S_i}{\partial E_i}\right)_{V_i, N_i=const}. \tag{11}$$

Equation 10 thus tells us that in equilibrium, $T_1 = T_2$. Note that this definition of temperature is completely general, and is *independent* on the internal properties of the molecules of any sub-system.

Notes:

[a]. The definition of temperature may look somewhat arbitrary. True, in Equation 11 we could replace $1/T_i$ with other functions of the temperature. However, as we will show later, this definition of temperature is identical with the perfect gas scale. Using Equations 9 and 11, we find that

$$\frac{dS_T}{dt} = \left(\frac{\partial S_T}{\partial E_1}\right) \left(\frac{dE_1}{dt}\right) = \left(\frac{1}{T_1} - \frac{1}{T_2}\right) \left(\frac{dE_1}{dt}\right) > 0. \tag{12}$$

The inequality sign represents the principle of increasing entropy.

It thus follows that if $T_1 < T_2$, then $dE_1/dt > 0$, namely heat flows from the subsystem with the higher temperature to that at lower temperature, in agreement with our everyday life experience.

[b]. In Equation 9, we used $(\partial S_T/\partial E_1)$, namely, we searched for the energy E_1 for which the total entropy is maximal. We *could not* search for the value of E_T for which the total entropy is maximal, since this energy is conserved in an isolated system.

[c]. As the energy E increases, the statistical weight $\Omega(E)$ also increases; this is because more energy in a system can be distributed in many more ways over the microscopic states of the system. This implies that the temperature defined in Equation 11 is **always positive**.

Example - 2.

Consider a system similar to the one considered in example 1, but now we allow the partition to move. Thus, the number of particles in each sub-system, N_1 and N_2 are fixed. However, the energies E_1 and E_2 can change (with the constraint $E_1 + E_2 = E_T$), and the volumes V_1 and V_2 can vary (with $V_1 + V_2 = V_T$).

As in the previous example, the equilibrium condition is that the total entropy of the system, S_T is maximized. In the equilibrium state, we expect both the temperature and the pressure on both sides to be equal.

Similar to the previous example, we chose E_1 and V_1 to be independent variables, and we search for the values that maximize S_T .

$$\begin{aligned} \left(\frac{\partial S_T}{\partial V_1}\right)_{V_T=const} &= \left(\frac{\partial S_1}{\partial V_1}\right) + \left(\frac{\partial S_2}{\partial V_1}\right) \\ &= \left(\frac{\partial S_1}{\partial V_1}\right) + \left(\frac{\partial S_2}{\partial V_2}\right) \left(\frac{\partial V_2}{\partial V_1}\right) \\ &= \left(\frac{\partial S_1}{\partial V_1}\right) - \left(\frac{\partial S_2}{\partial V_2}\right) = 0, \end{aligned} \tag{13}$$

or

$$\left(\frac{\partial S_1}{\partial V_1}\right)_{E_1, N_1} = \left(\frac{\partial S_2}{\partial V_2}\right)_{E_2, N_2}. \tag{14}$$

This must be interpreted as implying equal pressures in the two subsystems, since we know that in equilibrium $P_1 = P_2$. We thus *define* the pressure by

$$P_i \equiv T_i \left(\frac{\partial S_i}{\partial V_i}\right)_{E_i, N_i=const} \tag{15}$$

As we shall see, this definition of the pressure is identical to the conventional one, which appears, e.g., in the ideal gas equation of state, $PV = Nk_B T$. With this definition of the pressure, Equation 14 implies that in equilibrium, $P_1 = P_2$.

Similarly to the previous example, we can apply the principle of entropy increase to a case in which our system is not in equilibrium.

$$\begin{aligned} \frac{dS_T}{dt} &= \left(\frac{\partial S_T}{\partial V_1}\right) \left(\frac{dV_1}{dt}\right) \\ &= \left(\frac{\partial S_1}{\partial V_1} - \frac{\partial S_2}{\partial V_2}\right) \left(\frac{dV_1}{dt}\right) \\ &= \left(\frac{P_1}{T_1} - \frac{P_2}{T_2}\right) \left(\frac{dV_1}{dt}\right) > 0. \end{aligned} \tag{16}$$

In the case where $T_1 = T_2 = T$, if V_1 is increasing with time ($dV_1/dt > 0$), we obtain from equation 16, $(1/T)(P_1 - P_2) > 0$, which implies $P_1 > P_2$ - just as we would expect.

4. The Schottky defect

At absolute zero temperature, the atoms of a solid are ordered completely regularly on a crystal lattice. As the temperature increases, the atoms can change their positions: e.g., due to vibrations. As the atoms migrate, they leave vacant lattice sites, known as “point defects”. One kind of such defects are **Schottky defects**, in which the atoms migrate to the surface of the crystal (see Figure 4).

We now calculate the number of Schottky defects for a crystal in thermal equilibrium at some temperature T . Generally, atoms inside the lattice have lower energy than atoms at the surface, due to the stronger binding energy. Consider a lattice with N atoms and n Schottky defects, where we assume $n \ll N$, thus we can treat each defect separately. The energy associated with the defects is $E = n\epsilon$.

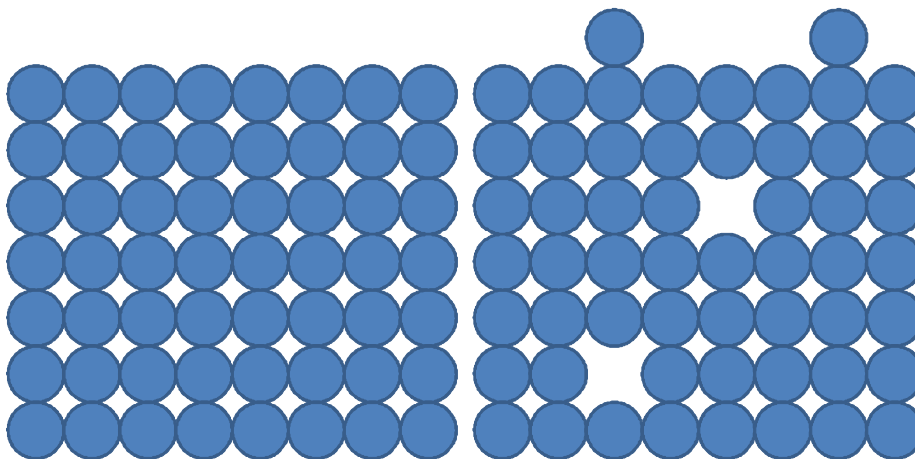


Fig. 4.— Schottky defect. Left: a two dimensional perfect lattice. Right: the same lattice with Schottky defects, namely two atoms migrated to the surface of the crystal leaving two vacancies.

In order to create n Schottky defects, we must move n atoms out of total N , thus we have

$$\Omega(n) = \binom{N}{n} = \frac{N!}{n!(N-n)!}, \quad (17)$$

and the associated entropy is

$$S(n) = k_B \ln \Omega(n) = k_B \ln \frac{N!}{n!(N-n)!} = k_B (\ln N! - \ln n! - \ln(N-n)!). \quad (18)$$

We can now use Stirling's formula,

$$\ln N! \approx N(\ln N - 1) \quad (19)$$

which is true for $N \gg 1$, to write

$$S(n) \approx k_B (N \ln N - n \ln n - (N-n) \ln(N-n)). \quad (20)$$

Furthermore, we know that the temperature is related to the entropy via

$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{dS(n)}{dn} \frac{dn}{dE} = \frac{1}{\epsilon} \frac{dS(n)}{dn} \quad (21)$$

Differentiating Equation 20 gives

$$\frac{dS(n)}{dn} = k_B (-\ln n + \ln(N-n)) = k_B \ln \left(\frac{N-n}{n} \right), \quad (22)$$

and thus

$$\frac{1}{T} = \frac{k_B}{\epsilon} \ln \left(\frac{N-n}{n} \right). \quad (23)$$

Obviously, it is much easier to measure T than n . We can thus use some algebra on Equation 23, to obtain

$$n = \frac{N}{\exp \frac{\epsilon}{k_B T} + 1}. \quad (24)$$

In the limiting case $\epsilon \gg k_B T$, $\exp(\epsilon/k_B T) \gg 1$ and $n \approx N e^{-\epsilon/k_B T}$ ($n \ll N$).

At $T = 0$ °K, $n = 0$ - all atoms are in place. This is the state of lowest energy, or *ground state*. At $T > 0$, $n > 0$. Typically, $\epsilon \sim 1$ eV (E.g., for Ge, $\epsilon \approx 1.7$ eV). Thus, at room temperature (290 °K, $k_B T \approx 1/40$ eV),

$$\frac{n}{N} = e^{-1.7/0.025} \approx 3 \times 10^{-30}. \quad (25)$$

However, at $T = 1000$ °K, $n/N \approx 3 \times 10^{-9}$ due to the exponential dependence.

5. Equilibrium of a system in a heat bath

So far we have considered the equilibrium of isolated systems. Now let us look at a system which is in contact with a heat bath having temperature T .

A **heat bath** is a body with a heat capacity very large compared to the system in question. This implies that the system and the heat bath can come to thermal equilibrium without a significant change in the temperature of the heat bath.

We can therefore look at this combined system as two sub-systems (the original system and the heat bath) which are in contact. However, since the heat bath is much larger than our system, its temperature is not changed, because of its size. The temperature of the original system, when in equilibrium, is also T , but its energy is not fixed. We assume that the combined system + heat bath are fully isolated from the rest of the world. **Let us see how to calculate the energy of the system.**

We assume that N and V are constants (no number change or volume change), and so the *macrostate* of the system in equilibrium is specified by T , V and N . The system has many different *microstates*, which we will denote $1, 2, \dots, r, \dots$. We will label their energies by $E_1, E_2, \dots, E_r, \dots$.

The energy by itself is not enough to fully determine the microstate. As we have seen in the paramagnet example, different microstates can have the same energy. Thus, we can only calculate the probability $P(r)$ of the system to be in microstate r having energy E_r .

In order to calculate $P(r)$, we recall that the combined system + heat bath are isolated from the rest of the world. Thus, the total energy of the combined system is constant - let us call it E_0 . If the system has energy E_r , then the energy of the heat bath is $E_0 - E_r$. Thus, the probability $P(r)$ must be proportional to the statistical weight of the heat bath to have energy $E_0 - E_r$, $\Omega_{HB}(E_0 - E_r)$,

$$P(r) \propto \Omega_{HB}(E_0 - E_r). \quad (26)$$

Since the probability of the system to be in a microstate is unity, the proportionality constant is obtained by summing over all possible microstates of the heat bath, namely

$$P(r) = \frac{\Omega_{HB}(E_0 - E_r)}{\sum_r \Omega_{HB}(E_0 - E_r)}. \quad (27)$$

We can now use the definition of entropy, $S = k_B \ln \Omega$, to write $\Omega = e^{S/k_B}$. We can thus write the probability by

$$P(r) = \frac{e^{\frac{S_{HB}(E_0 - E_r)}{k_B}}}{\sum_r e^{\frac{S_{HB}(E_0 - E_r)}{k_B}}}. \quad (28)$$

Equation 28 is completely general, and is true for any system composed of two sub-systems. Now, we make use of the assumption that one of the sub-systems is a heat bath, namely its energy is much larger compared to that of the system: $E_r \ll E_0$. This assumption enables us to expand $S_{HB}(E_0 - E_r)$ into a Taylor series:

$$S_{HB}(E_0 - E_r) = S_{HB}(E_0) - E_r \frac{\partial S_{HB}(E_0)}{\partial E_0} + \frac{E_r^2}{2} \frac{\partial^2 S_{HB}(E_0)}{\partial E_0^2} + \dots \quad (29)$$

Using the definition of the temperature:

$$\frac{\partial S_{HB}(E_0)}{\partial E_0} = \frac{1}{T}, \quad (30)$$

and

$$\frac{\partial^2 S_{HB}(E_0)}{\partial E_0^2} = \frac{\partial}{\partial E_0} \left(\frac{\partial S_{HB}(E_0)}{\partial E_0} \right) = \frac{\partial}{\partial E_0} \left(\frac{1}{T} \right) = 0, \quad (31)$$

hence we can write Equation 29 as

$$S_{HB}(E_0 - E_r) = S_{HB}(E_0) - \frac{E_r}{T}, \quad (32)$$

or

$$\frac{S_{HB}(E_0 - E_r)}{k_B} = \frac{S_{HB}(E_0)}{k_B} - \beta E_r, \quad (33)$$

where

$$\boxed{\beta \equiv \frac{1}{k_B T}} \quad (34)$$

is known as the **temperature parameter**, and naturally occurs in many statistical mechanics equations.

Using this information we can now write the probability $P(r)$ in Equation 28 in the form

$$P(r) = \frac{e^{\frac{S_{HB}(E_0 - E_r)}{k_B}}}{\sum_r e^{\frac{S_{HB}(E_0 - E_r)}{k_B}}} = \frac{e^{\frac{S_{HB}(E_0)}{k_B} - \beta E_r}}{\sum_r e^{\frac{S_{HB}(E_0)}{k_B} - \beta E_r}} = \frac{e^{-\beta E_r}}{\sum_r e^{-\beta E_r}} = \frac{1}{Z} e^{-\beta E_r}. \quad (35)$$

where

$$\boxed{Z \equiv \sum_r e^{-\beta E_r}} \quad (36)$$

is called the **partition function**. It plays a central role in studying the properties of systems at a fixed temperature.

Equation 35 is known as **Boltzmann distribution**. It gives the *probability that a system, when placed in a heat bath at temperature T be in the microstate r with energy E_r* . We see that this probability depends on the energy E_r of the state r . The *only* property of the heat bath on which it depends is its temperature, T .

5.1. The partition function and the mean energy

The partition function in Equation 36 is a sum over all microstates of the system. As we noted in the paramagnet example, different microstates may have the same energy. We can therefore re-write the partition function in Equation 36 as a sum over all *energies*:

$$Z \equiv \sum_r e^{-\beta E_r} = \sum_{E_r} g(E_r) e^{-\beta E_r}. \quad (37)$$

In Equation 37, the summation is over all different **energies**, E_r , and $g(E_r)$ is the number of microstates all having the same energy E_r . The number $g(E_r)$ is known as the **degeneracy** of the energy E_r .

Using the same line of reasoning, we can write the probability $P(E_r)$ of the system to be in a state with energy E_r ,

$$P(E_r) = g(E_r)P(r) = \frac{1}{Z}g(E_r)e^{-\beta E_r}. \quad (38)$$

From Boltzmann distribution we can obtain the mean energy \bar{E} of the system in contact with a heat bath:

$$\bar{E} = \sum_r P(r)E_r = \frac{1}{Z} \sum_r E_r e^{-\beta E_r} \quad (39)$$

Using the definition of Z from Equation 36, we find that $\partial Z/\partial\beta = \sum_r (-E_r)e^{-\beta E_r}$, which is just the summation in Equation 39. We can thus re-write Equation 39 as

$$\boxed{\bar{E} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta}.} \quad (40)$$

Thus, we obtained the average energy of the system in terms of the partition function.

Note that we used partial derivatives as the energy levels E_r are held constant: these energies *do not* depend on the temperature, but purely on the macroscopic structure of the system. (Recall that in the paramagnet example, the energy levels depend on the external magnetic field).

5.2. Fluctuations around the mean energy

Equation 40 gives the *mean* energy of the system. However, its *actual* energy will fluctuate, as it is in contact with the heat bath. How large are these fluctuations ?

The magnitude of the fluctuations is measured by the **standard deviation**, ΔE , defined by

$$(\Delta E)^2 \equiv \overline{(E - \bar{E})^2} = \overline{E^2} - \bar{E}^2 \quad (41)$$

In order to calculate ΔE^2 , we calculate

$$\begin{aligned} \frac{\partial^2 \ln Z}{\partial \beta^2} &= \frac{\partial}{\partial \beta} \left(\frac{\partial \ln Z}{\partial \beta} \right) \\ &= \frac{\partial}{\partial \beta} \left(-\frac{1}{Z} \sum_r E_r e^{-\beta E_r} \right) \\ &= \frac{1}{Z^2} \frac{\partial Z}{\partial \beta} \sum_r E_r e^{-\beta E_r} + \frac{1}{Z} \sum_r E_r^2 e^{-\beta E_r} \\ &= -\bar{E}^2 + \sum_r E_r^2 P(r) \end{aligned} \quad (42)$$

Since the average of a function $f(E_r)$ is $\overline{f(E_r)} \equiv \sum_r f(E_r)P(r)$, the last term in Equation 42 is just $\overline{E^2}$. We thus obtain

$$(\Delta E)^2 = \frac{\partial^2 \ln Z}{\partial \beta^2} = -\frac{\partial \bar{E}}{\partial \beta} = -\frac{dT}{d\beta} \frac{\partial \bar{E}}{\partial T} = k_B T^2 C, \quad (43)$$

where $C \equiv \partial \bar{E} / \partial T$ is the heat capacity of the system at constant external parameters.

The magnitude of the fluctuations is therefore given by

$$\frac{\Delta E}{\bar{E}} = \frac{(k_B T^2 C)^{1/2}}{\bar{E}} \quad (44)$$

The crucial point is that both C and \bar{E} are **extensive** parameters, namely they are proportional to the number of atoms in the system, N ; while $k_B T^2$ is independent of N . Therefore, the dependence of $\Delta E / \bar{E}$ on the size of the system is

$$\frac{\Delta E}{\bar{E}} \sim \frac{1}{\sqrt{N}}. \quad (45)$$

Thus, for macroscopic systems with $N \sim 10^{23}$, $\Delta E / \bar{E} \sim 10^{-11}$!. This means that *the fluctuations are extremely small; the energy of a macroscopic body in a heat bath is completely determined, for any practical purpose*. This is why we can use statistical physics to determine quantitatively the properties of macroscopic systems: the relative fluctuations are always of the order $N^{-1/2}$, which is tiny.

The fact that $\Delta E / \bar{E}$ is so small implies that the probability distribution (Equation 38) has an extremely sharp maximum at energy \bar{E} . This is true for $N \gg 1$; in the limit of small N , this no longer holds.

6. Extending the definition of entropy

Previously, we defined entropy for *isolated* systems. Now, let us generalize this definition, so that it would be applicable also for systems in contact with a heat bath.

Consider a general macroscopic system. Let us label its microstates by $1, 2, \dots, r, \dots$. Let $P(r)$ be the probability that the system is in the microstate r . At this point, we know nothing about $P(r)$, except for the fact that the system must be at **a** state, hence

$$\sum_r P(r) = 1. \quad (46)$$

What is the entropy of the system ? In order to answer this question, let us consider an ensemble of very large number, say ν of identical replicas of our system. Thus, all these systems are in thermal equilibrium with the heat bath. We can now ask: “how many of these systems are in the microstate r ?” - the answer is

$$\nu_r = \nu P(r). \quad (47)$$

The statistical weight Ω_ν of the ensemble in which ν_1 of the systems are in state 1, ν_2 are in state 2,... is the number of ways in which this particular distribution can be realized (same logic as in the paramagnetic solid example), which is

$$\Omega_\nu = \frac{\nu!}{\nu_1! \nu_2! \dots \nu_r! \dots} \quad (48)$$

We can now use Boltzmann’s definition of entropy, to write

$$\begin{aligned} S_\nu &= k_B \ln \Omega_\nu \\ &= k_B \ln \left(\frac{\nu!}{\nu_1! \nu_2! \dots \nu_r! \dots} \right) \\ &= k_B (\ln \nu! - \sum_r \ln \nu_r!) \\ &= k_B (\nu \ln \nu - \nu - \sum_r \nu_r \ln \nu_r + \sum_r \nu_r) \\ &= k_B (\nu \ln \nu - \sum_r \nu_r \ln \nu_r), \end{aligned} \quad (49)$$

where we have used Stirling’s formula.

Using now $\nu_r = \nu P(r)$, we get

$$\begin{aligned} S_\nu &= k_B (\nu \ln \nu - \sum_r \nu P(r) \ln (\nu P(r))) \\ &= k_B (\nu \ln \nu - \nu \ln \nu \sum_r P(r) - \nu \sum_r P(r) \ln P(r)) \\ &= -k_B \nu \sum_r P(r) \ln P(r) \end{aligned} \quad (50)$$

where in the last line we used the normalization, $\sum_r P(r) = 1$.

This entropy is the entropy of the total ensemble of systems; since entropy is an additive quantity, the entropy of a single system is given by

$$S = \frac{1}{\nu} S_\nu = -k_B \sum_r P(r) \ln P(r). \quad (51)$$

We can now calculate the entropy of a system in contact with a heat bath with temperature T : using the probability $P(r)$ from Equation 35, we find

$$\begin{aligned} S &= -k_B \sum_r \left(\frac{e^{-\beta E_r}}{Z} \right) \ln \left(\frac{e^{-\beta E_r}}{Z} \right) \\ &= k_B \sum_r \left(\frac{e^{-\beta E_r}}{Z} \right) (\beta E_r + \ln Z) \end{aligned} \quad (52)$$

We can use now the definition of Z from Equation 36 and \bar{E} from Equation 39 to write

$$\begin{aligned} S &= k_B \beta \bar{E} + k_B \sum_r \frac{e^{-\beta E_r}}{Z} \ln Z \\ &= \frac{\bar{E}}{T} + k_B \ln Z \end{aligned} \quad (53)$$

Note that in the new definition of the entropy in Equation 53, the entropy is a function of T , V and N , while in the definition for isolated systems, $S = k_B \ln \Omega(E, V, N)$ is a function of E , V and N . However, as we saw, for macroscopic systems at temperature T , the energy fluctuations are tiny. Therefore the energy is well defined by its average, \bar{E} .

7. Helmholtz free energy

When we discussed isolated systems, the basic statistical quantity was the statistical weight $\Omega(E, V, N)$, and the basic thermodynamic quantity was the entropy, $S(E, V, N)$. When considering a system in contact with a heat bath the basic statistical quantity is the partition function, $Z(T, V, N)$. Thus, we can find the corresponding thermodynamics quantity, which is the **Helmholtz free energy**, defined by

$$\boxed{F(T, V, N) \equiv -k_B T \ln Z(T, V, N)} \quad (54)$$

Using Equation 53, we can write

$$\begin{aligned} S &= \frac{\bar{E}}{T} - \frac{F}{T}, \\ F &= \bar{E} - TS \end{aligned} \quad (55)$$

In fact, in macroscopic systems we can safely replace \bar{E} with E , as the fluctuations are so tiny.

Thus, for a system in equilibrium in contact with heat bath and having a constant volume, the Helmholtz free energy plays a similar role to that of entropy in an isolated system. In particular, just as the entropy S is maximal in equilibrium in an isolated system, so the Helmholtz free energy F obtains its minimum in equilibrium state of a system having constant volume in a heat bath (namely, T , V and N are constants).

7.1. Example: Helmholtz free energy and the Schottky defect

Let us return to the example of Schottky defect introduced earlier. Recall that we had a solid crystal with N atoms and n defects. The energy per defect is ϵ , and so the total energy is $E = n\epsilon$. The entropy was calculated in Equation 20 to be

$$S(n) \approx k_B (N \ln N - n \ln n - (N - n) \ln(N - n)), \quad (56)$$

and thus the Helmholtz free energy is

$$F = E - TS = n\epsilon - k_B T (N \ln N - n \ln n - (N - n) \ln(N - n)). \quad (57)$$

We know that F obtains its minimum value in Equilibrium state, so we need to find dF/dn and equate to 0,

$$\begin{aligned} \frac{dF}{dn} = \epsilon - T \frac{dS}{dn} &= \epsilon - k_B T (-\ln n + \ln(N - n)) \\ &= \epsilon - k_B T \ln \left(\frac{N-n}{n} \right) = 0. \end{aligned} \quad (58)$$

The solution is

$$n = N \left(e^{\epsilon/k_B T} + 1 \right)^{-1}. \quad (59)$$

For $\epsilon \gg k_B T$,

$$n \approx N e^{-\epsilon/k_B T}. \quad (60)$$

These are the same results found when maximizing the entropy (see Equation 24).