Physics 112

The (generalized) free energy is a minimum in equilibrium

Peter Young

I. INTRODUCTION

When you understand *well* the contents of this handout, you will have understood *most* of the formalism of equilibrium statistical mechanics needed for this course.

(The remaining piece of formalism concerns the "chemical potential" which we will discuss later.)

II. PROOF THAT THE GENERALIZED FREE ENERGY IS A MINIMUM IN EQUILIBRIUM

Consider a system in thermal contact with a reservoir at temperature T. Its free energy F is related to its energy U and entropy $S \equiv k_B \, \sigma$, by

$$F = U - TS. \tag{1}$$

We recall that it is also related to the partition function Z, and hence to the energy levels of the system E_n, by

$$\boxed{F = -k_B T \ln Z = -k_B T \ln \left(\sum_n e^{-\beta E_n} \right),} \tag{2}$$

where $\beta = 1/k_B T$. It is convenient to separate the sum over all states into (i) a sum over all states with a fixed energy U, which just gives $g(U)e^{-\beta U}$ where $g(U)$ is the multiplicity factor (number of states with energy U), and (ii) a sum over different values of the energy, *i.e.*

$$F = -k_B T \ln \left(\sum_U g(U) e^{-\beta U} \right). \tag{3}$$

Writing

$$\boxed{g(U) = e^{\widetilde{S}(U)/k_B},} \tag{4}$$

where $\widetilde{S}(U)$ is the entropy if the system is constrained to have energy U, we can rewrite this last equation as

$$F = -k_B T \ln \left(\sum_U e^{-\beta [U - T\widetilde{S}(U)]} \right) = -k_B T \ln \left(\sum_U e^{-\beta \widetilde{F}(U)} \right), \tag{5}$$

where

$$\widetilde{F}(U) = U - T\widetilde{S}(U) \tag{6}$$

is the free energy of the system *constrained to have energy U*. We put a "tilde" on $F(U)$ and $S(U)$ to indicate that they are not, in general, the equilibrium free energy and entropy, but describe *non-equilibrium* quantities. In a previous edition of the book, which was just by Kittel, the terms "generalized free energy" and "generalized entropy" were used for such non-equilibrium objects, and I find this to be helpful in distinguishing them from the corresponding equilibrium values. I will therefore use this terminology here.

Consider the value of U for which $e^{-\beta\widetilde{F}(U)}$ is a *maximum, i.e.* for which $\widetilde{F}(U)$ is a *minimum*. Let us call this value $U^\star$. Now $\widetilde{F}(U)$ is extensive (*i.e.* proportional to N where N is generally very large) and so, as we have discussed earlier in the context of the entropy, $e^{-\beta\widetilde{F}(U)}$ is *so much* bigger at $U = U^\star$ than at any other value of U, that this value completely dominates the sum. From Eq. (5), this means we have

$$-\beta F \equiv \ln\left(\sum_U e^{-\beta\widetilde{F}(U)}\right) \simeq \ln\left(e^{-\beta\widetilde{F}(U^\star)}\right) = -\beta\widetilde{F}(U^\star), \tag{7}$$

with an error which is negligible (of order $\ln(N)/N$ for large N). In other words,

$$F = \widetilde{F}(U^\star), \tag{8}$$

so $\boxed{\widetilde{F}(U^\star), \text{ the } minimum \text{ of } \widetilde{F}(U), \text{ is the equilibrium value of the free energy.}}$

According to the Boltzmann distribution, $\exp(-\beta U)$ is proportional to the probability that the system is in a *single* state with energy U, so it follows that $P(U)$, the probability that the system is *any* state with energy U, is proportional to $g(U)\,e^{-\beta U}$ which, from Eq. (4), is just $e^{-\beta\widetilde{F}(U)}$, *i.e.*

$$P(U) \propto e^{-\beta\widetilde{F}(U)}. \tag{9}$$

Now suppose that the system initially has an energy different from the equilibrium value. Then, according to Eq. (9), the most probable result is that the energy will change in such a way as to minimize the generalized free energy. Eventually the generalized free energy will have reached its minimum (*i.e.* most probable) value, and will then become independent of time. In this equilibrium situation, in which the energy is $U^\star$, the generalized free energy $\widetilde{F}(U^\star)$ will equal the equilibrium free energy F, see Eq. (8).

Note that the quantity here which is allowed to vary is the energy, while other constraints (such as pressure, magnetic field etc.) are kept constant. One has to be careful to understand what is allowed to vary and what is being kept constant.

III. DIRECT (EQUILIBRIUM) CALCULATION OF THE FREE ENERGY FOR THE TOY MODEL

Let us work out $\widetilde{F}$ for the toy model with N spins in a magnetic field B. In this handout we work in units where the magnetic moment of one spin, μ, is equal to unity. The energy is given by

$$E = -B \sum_{i=1}^{N} \mu_i = -MB \,, \tag{10}$$

where $\mu_i \pm 1$. We call $M = \sum_i \mu_i$ the magnetization.

It is easy to get the equilibrium value of M since each spin is independent of the others so $\langle \mu_i \rangle$ can be obtained by just considering the states of the single spin i. The two possibilities are:

$$\mu_i = 1, \ \ E = -B \tag{11}$$

$$\mu_i = -1, \ \ E = B \,. \tag{12}$$

Hence the partition function for a single spin is

$$\boxed{z = e^{B/k_B T} + e^{-B/k_B T} = 2\cosh(B/k_B T) \,,} \tag{13}$$

and the probabilities of the $\mu_i = 1$ and $\mu_i = -1$ states are

$$P_1 = \frac{e^{B/k_B T}}{z}, \qquad P_{-1} = \frac{e^{-B/k_B T}}{z} \,. \tag{14}$$

It follows that $\langle \mu_i \rangle$, the average magnetization of site i, which we will write as m, is given by

$$m \equiv \langle \mu_i \rangle \ = \ 1\,P_1 + (-1)\,P_{-1} = \frac{e^{B/k_B T} - e^{-B/k_B T}}{e^{B/k_B T} + e^{-B/k_B T}}$$

$$= \ \boxed{\tanh\left(\frac{B}{k_B T}\right) \,.} \tag{15}$$

The *total* magnetization of the N spins is just

$$M = mN \,. \tag{16}$$

Note that it is immaterial whether we use the magnetization M of a single configuration, or the average magnetization $\langle M \rangle$, in Eq. (16) because, for large N, the probability of getting M significantly different from $\langle M \rangle$ is negligible.

We can also determine the free energy from Eq. (2). For one spin we have

$$f = -k_B T \ln z = \boxed{-k_B T \ln\left[2\cosh\left(\frac{B}{k_B T}\right)\right] \,.} \tag{17}$$

Now since we sum over all configurations of the spin independently, the partition function of the whole system, Z, is just the product of the partition functions of each spin, *i.e.*

$$Z = z^N. \tag{18}$$

Hence the total free energy F is just N times the free energy of a single spin, f, *i.e.*

$$F = -Nk_BT \ln\left[2\cosh\left(\frac{B}{k_BT}\right)\right]. \tag{19}$$

IV. FREE ENERGY OF THE TOY MODEL OBTAINED BY MINIMIZING THE GENERALIZED FREE ENERGY

What we just discussed is the most *convenient* way to obtain m; compute the the partition function (an equilibrium quantity) and, from that, obtain the probabilities of the various states being occupied. However, we will now verify that we do get the same result by going through the more long-winded approach of determining $\widetilde{F}(U)$ and minimizing it.

Writing the energy as U rather than E we have, from Eqs. (10) and (16),

$$U = -MB = -NmB, \tag{20}$$

and so minimizing $\widetilde{F}$ with respect to U is equivalent to minimizing it with respect to M or m. The log of the multiplicity function $\ln g(N, M)$ (where $M = 2k$ with k the spin-imbalance factor in the book) is given by

$$\ln g(N, M) = \ln N! - \ln((N-M)/2)! - \ln((N+M)/2)!. \tag{21}$$

Using Stirling's approximation $\ln N! \simeq N \ln N - N$ gives

$$\ln g(N, M) = N \ln N - \frac{N-M}{2}\ln\left(\frac{N-M}{2}\right) - \frac{N+M}{2}\ln\left(\frac{N+M}{2}\right). \tag{22}$$

Writing $N \ln N = ((N-M)/2)\ln N + ((N+M)/2)\ln N$ we can write Eq. (22) as

$$\ln g(N, M) = -N\left[\frac{1+m}{2}\ln\left(\frac{1+m}{2}\right) + \frac{1-m}{2}\ln\left(\frac{1-m}{2}\right)\right], \tag{23}$$

where $m = M/N$, see Eq. (16).

Hence, from Eqs. (4), (6), (20), and (23), the generalized free energy (which later on in the course we will call the "Landau function" in a similar context) is given by

$$\boxed{\widetilde{f}(m) \equiv \frac{\widetilde{F}(m)}{N} = -mB + k_BT\left[\frac{1+m}{2}\ln\left(\frac{1+m}{2}\right) + \frac{1-m}{2}\ln\left(\frac{1-m}{2}\right)\right],} \tag{24}$$

in which the first term on the RHS, proportional to B, is the energy and the second, proportional to T, is $-T$ times the (generalized) entropy, see Eq. (6). Note that the expression for the entropy is that expected from the Boltzmann definition, discussed in class, $s/k_B = -\sum_l P_l \ln P_l$, in which, for the single spin situation here, l runs over the two states "↑" and "↓", with $P_\uparrow = (1+m)/2$ and $P_\downarrow = (1-m)/2$, *i.e.*

$$\frac{\widetilde{s}(m)}{k_B} = \frac{\widetilde{S}(m)}{k_B N} = -\left[\frac{1+m}{2}\ln\left(\frac{1+m}{2}\right) + \frac{1-m}{2}\ln\left(\frac{1-m}{2}\right)\right] = -\left[P_\uparrow \ln P_\uparrow + P_\downarrow \ln P_\downarrow\right]. \quad (25)$$

Note too that $\widetilde{F}(m)$ is extensive (*i.e.* proportional to N) as expected.

Minimizing $\widetilde{F}(m)$ with respect to m, and calling the value at the minimum $m^\star$ gives

$$B = \frac{k_B T}{2}\ln\left(\frac{1+m^\star}{1-m^\star}\right), \quad (26)$$

which is equivalent to

$$\frac{1+m^\star}{1-m^\star} = e^{2B/k_B T}. \quad (27)$$

Rearranging gives

$$m^\star = \frac{e^{2B/k_B T}-1}{e^{2B/k_B T}+1} = \boxed{\tanh\left(\frac{B}{k_B T}\right)}, \quad (28)$$

which agrees with Eq. (15), which was obtained (much more easily) from the partition function. Similarly one can show, after a bit of algebra, that substituting $m^\star$ from Eq. (28) into Eq. (24) gives

$$\widetilde{F}(m^\star) = F, \quad (29)$$

where F is given by Eq. (19).

As an example, the figure below plots the generalized free energy per site, $\widetilde{f}(m) = \widetilde{F}(m)/N$ from Eq. (24) for $B = T = 1$ (in the figures we use units where $k_B = 1$).
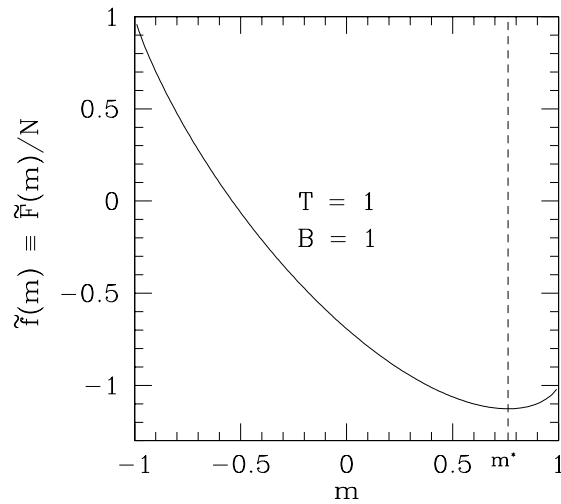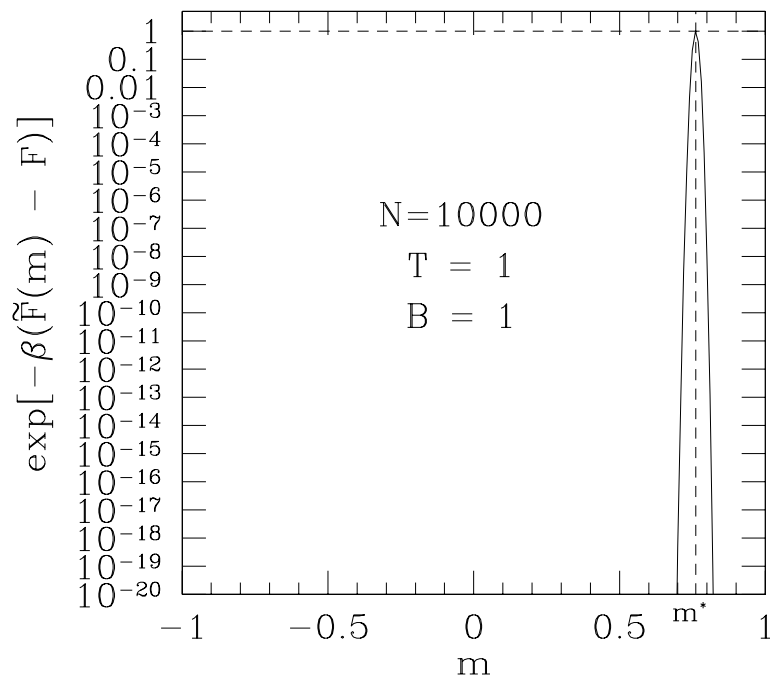
The curve is smooth and has a minimum at $m^\star = \tanh(B/T) = \tanh 1 = 0.7616$.

Although $\widetilde{F}(m)/N$ has a *shallow* minimum, the probability of a given m, which is proportional to $\exp\left[-\beta\widetilde{F}(m)\right]$, has a very *sharp* peak for large N. The figure below plots

$$\frac{e^{-\beta\widetilde{F}(m)}}{e^{-\beta F}}, \tag{30}$$

where F is the equilibrium free energy given in Eq. (19) and $\widetilde{F}(m)$ is given in Eq. (24), for $N = 10^4, T = B = 1$.

The peak is indeed seen to be incredibly sharp, especially if one notes that the vertical axis is logarithmic.

Equation (30) describes an individual term in Eq. (5) divided by the largest term. The denominator of Eq. (30) *is* the largest term because $\widetilde{F}(m^\star) = F$ (with negligible error for large N) as shown in Eq. (8). The largest value of Eq. (30) is therefore unity by definition. The figure therefore represents the (unnormalized) probability of getting different values of m, see also Eq. (9). Remember that the data is for $N = 10000$. One clearly sees a *very sharp peak* even though I used an extended logarithmic vertical scale. The equilibrium value of m is $m^\star = 0.7616$, which is shown by the dashed vertical line. This is at the peak as expected. This figure clearly shows that

the probability of getting a value of m significantly different from the equilibrium value is negligible

for a large system.

I emphasize that the *shallow* minimum in $\widetilde{f}(m) = \widetilde{F}(m)/N$ in the first figure becomes the *sharp* maximum in the second figure because $\widetilde{f}(m)$ is multiplied by N (which is large) *and* put in an exponential. (The reason that the minimum becomes a maximum is trivial: the minus sign in the exponents in Eq. (30).)

V.   CONCLUSION

In this handout we have seen that the (generalized) free energy of a system in thermal contact with a reservoir is a minimum in equilibrium. Furthermore, for a large system, the probability that system deviates significantly from its equilibrium state is negligible.

In addition, we showed that the simplest way to determine the equilibrium value of m of our toy model is to use the Boltzmann distribution and compute the partition function. Using the Boltzmann distribution the system can exchange energy with a much larger system (reservoir). It is easier to allow these energy exchanges than to work with a closed system (in which the energy is fixed) because, for a closed system, the resulting constraint on the total energy make the calculations cumbersome. (One needs to calculate the multiplicity factor of the whole system for a fixed energy.)

Quite generally, the easiest way to determine equilibrium properties in statistical mechanics is to compute the partition function and Boltzmann probabilities. Later in the course, when we explicitly consider identical quantum particles (bosons and fermions) we will find it convenient to allow the system to exchange not only energy *but also particles* with the reservoir. (This change is necessary because indistinguishability of the particles changes the way we count states.) Doing so

will lead to a straightforward extension of the definition of the partition function to include a sum over particle number, as well as a sum over states (for each particle number). This sum is then called the "*grand partition function*", the log of which will give a free-energy type quantity, called the "grand potential".