The main formulae in statistical thermal physics

Andrew M. Steane
Exeter College and Department of Atomic and Laser Physics, University of Oxford.

January 31, 2026

Abstract

Derivations of formulae for internal energy, entropy, temperature and free energy.

1 Some basic formulae

In this section we consider only thermal equilibrium, so that the Boltzmann distribution applies:

$$p_i = \frac{1}{Z} e^{-\beta \epsilon_i} \tag{1}$$

where i labels individual quantum states of the system in question, and Z is the partition function **defined** by

Partition function

$$Z = \sum_i e^{-\beta \epsilon_i} \tag{2}$$

where the sum is over all the quantum states of the system[1], ϵ_i is the energy of state i and β is a parameter which we will show shortly is related to temperature.

Eqn (2) can also be written $Z = \mathrm{Tr}(e^{-\beta H})$ where $\mathrm{Tr}(\cdots)$ signifies a trace operation, and H is the Hamiltonian written in the form of a matrix. In the basis of energy eigenstates, H is diagonal with the eigenvalues ϵ_i on the diagonal, so the trace expression reproduces (2). (The reader unfamiliar with matrix methods can ignore this and simply use (2)).

We will now show how various thermodynamic quantities are related to the partition function.

First consider the average energy which is given by

$$\bar{U} = \sum_i p_i \epsilon_i. \tag{3}$$

To find this, observe that

$$\frac{\partial Z}{\partial \beta} = \sum_i -\epsilon_i e^{-\beta \epsilon_i} = -\sum_i \epsilon_i (Z p_i) \tag{4}$$

where in the partial derivative all the energies ϵ_i are constant. It follows that

---

[1] To be precise we mean the mutually orthogonal energy eigenstates.

Internal energy

$$\bar{U} = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta}. \tag{5}$$

Next we find the entropy:

$$\begin{aligned} S &= -k_{\mathrm{B}} \sum_i p_i \ln p_i \;=\; k_{\mathrm{B}} \sum_i p_i (\beta \epsilon_i + \ln Z) \\ &= k_{\mathrm{B}} \left(\beta \bar{U} + \ln Z \right) \end{aligned} \tag{6}$$

Hence

Entropy, version 1

$$S = k_{\mathrm{B}} \left(-\beta \frac{\partial \ln Z}{\partial \beta} + \ln Z \right). \tag{7}$$

Next we shall find the temperature. The fundamental relation of thermodynamics (1st and 2nd law) is

$$\mathrm{d}U = T\mathrm{d}S + X\mathrm{d}R \tag{8}$$

where X and R are suitable macroscopic variables such that $X\mathrm{d}R$ expresses the work in a reversible change in a simple system.[2] Hence

$$\frac{1}{T} = \left.\frac{\partial S}{\partial U}\right|_R \tag{9}$$

We have

$$\left.\frac{\partial S}{\partial \beta}\right|_{\epsilon} = k_{\mathrm{B}} \left(\left.\frac{\partial \ln Z}{\partial \beta}\right|_{\epsilon} + \bar{U} + \beta \left.\frac{\partial \bar{U}}{\partial \beta}\right|_{\epsilon} \right) = k_{\mathrm{B}} \beta \left.\frac{\partial \bar{U}}{\partial \beta}\right|_{\epsilon} \tag{10}$$

where the constant ϵ is a shorthand notation for the condition that all the energy levels are constant. It follows that

$$\left.\frac{\partial S}{\partial \bar{U}}\right|_{\epsilon} = k_{\mathrm{B}} \beta. \tag{11}$$

We now argue that constant energy levels ϵ_i corresponds to constant R in the thermodynamic expression. Therefore

Temperature

$$T = \frac{1}{k_{\mathrm{B}} \beta}, \qquad \text{or} \quad \beta = \frac{1}{k_{\mathrm{B}} T} \tag{12}$$

where we have taken $\bar{U}$ as the internal energy in the thermodynamic expression.

For a system with internal energy U the *Helmholtz function* F is defined by

$$F = U - TS. \tag{13}$$

Using (5), (6) and (12) we have $\bar{F} = \bar{U} - k_{\mathrm{B}} T (\beta \bar{U} + \ln Z)$ which gives

Helmholtz function

$$\bar{F} = -k_{\mathrm{B}} T \ln Z. \tag{14}$$

---

[2] More generally there could be many pairs of such variables giving rise to further terms in the expression for $\mathrm{d}U$; this does not affect the argument here if we allow ‘constant R’ to stand in for ‘constant all such quantities’ so that (9) still holds.

This relationship between Z and $\bar{F}$ can also be expressed $Z = \exp(-\beta \bar{F})$ which is reminiscent of the Boltzmann distribution and can serve as a mnemonic for this important result.

Because (14) is easy to remember, it can serve as a useful alternative way to get the entropy (i.e. an alternative to (5)). We recall the thermodynamic relationship

$$\mathrm{d}F = -S\mathrm{d}T + X\mathrm{d}R \tag{15}$$

(which follows immediately from the definition $F = U - TS$ and the fundamental thermodynamic relation (8)), and thus obtain

$$S = -\left.\frac{\partial F}{\partial T}\right|_R . \tag{16}$$

In other words:

Entropy, more memorable version

$$S = -\frac{\partial F}{\partial T} = \frac{\partial}{\partial T}\left(k_{\mathrm{B}} T \ln Z\right) \tag{17}$$

1.1 'Single particle' partition function

The expressions (1)–(14) all apply to a single system, of whatever size and composition, in thermal equilibrium. In practice for many systems it may be difficult to analyse the energy level structure in full, and difficult to obtain the partition function. An important simplification occurs if our system is itself an ensemble of many smaller, simpler systems which are only weakly interacting. In this situation the partition function of one of the small parts is written Z_1 and if the systems are distinguishable one finds

$$Z = Z_1^N \tag{18}$$

(see exercise 2). Notice that all the main macroscopic quantities mentioned in (5)–(14) depend on $\ln Z$ not Z directly, so they can all be obtained from $N \ln Z_1$ and indeed one finds

$$\bar{U} = N\bar{U}_1, \qquad S = NS_1, \qquad \bar{F} = N\bar{F}_1 \tag{19}$$

in an obvious notation. For systems of this type the quantity Z_1 is often called the 'single particle partition function' even though the small parts do not need to be single particles; they can be anything at all as long as the system is a weakly-interacting set of them.

A small confusion that can arise is that some texts seem to imply that when we treat Z_1 we are treating a microcanonical ensemble. In fact we are not, since the definition of Z_1 implies that the probabilities follow the Boltzmann distribution, so we are treating thermal equilibrium, which is equivalent to a canonical ensemble. The confusion arises because there is a finite number N of subsystems. But the entire collection of N subsystems can itself be in equilibrium with a thermal reservoir, so it is perfectly possible to have precisely thermal equilibrium conditions at a well-defined temperature for a finite system, and this is what we have typically in mind when we employ the single-particle partition function.

1.2 Heat capacity

The heat capacity plays an important role because it is often the most directly measurable physical quantity from which information about energy and entropy can be obtained. A heat capacity always has the form

$$\frac{đQ_x}{\mathrm{d}T} = T\left.\frac{\partial S}{\partial T}\right|_x \tag{20}$$

where Q is heat and T is temperature, but a given system has an infinite number of different heat capacities owing to the fact that $đQ$ depends on the type of change the system undergoes when it is being supplied with heat. For a fluid the two most important heat capacities (called *principle heat capacities*) are the constant-pressure and constant-volume heat capacities, written C_p and C_V. In statistical physics the most important heat capacity, for any system, is the one describing heating or cooling such that the energy eigenstates and their energies do not change. For this heat capacity, and only this one, the change in internal energy is equal to the heat entering, and therefore

heat capacity

$$C = T\left.\frac{\partial S}{\partial T}\right|_{\epsilon} = \left.\frac{\partial \bar{U}}{\partial T}\right|_{\epsilon} \tag{21}$$

where the ϵ indicates constant energy levels. This 'constant ϵ' is taken as understood when the expression is written without explicitly indicating what is constant in the partial derivative.

There is an interesting relationship between this heat capacity and the system's energy fluctuations in thermal equilibrium. Starting from the expression for mean energy,

$$\bar{U} = \sum_i p_i \epsilon_i = \sum_i \frac{e^{-\beta\epsilon_i}}{Z}\epsilon_i \tag{22}$$

one obtains

$$\begin{aligned}\frac{\partial \bar{U}}{\partial \beta} &= \sum_i -\frac{e^{-\beta\epsilon_i}}{Z}\epsilon_i^2 - \frac{1}{Z^2}e^{-\beta\epsilon_i}\frac{\partial Z}{\partial \beta} \\ &= -\langle \epsilon^2 \rangle - \sum_i p_i\epsilon_i \frac{1}{Z}\frac{\partial Z}{\partial \beta} \\ &= -\langle U^2 \rangle + \langle U \rangle^2 \end{aligned} \tag{23}$$

where in the last step we used that $\langle U^2 \rangle$ is simply another name for $\langle \epsilon^2 \rangle$ (the symbols U and ϵ refer to the same quantity), and we used the expression (5) giving $\bar{U}$ in terms of Z. We have, therefore,

$$\frac{\partial \bar{U}}{\partial \beta} = -\Delta U^2 \tag{24}$$

where ΔU^2 is the variance of the energy distribution. But since

$$C = \frac{\partial \bar{U}}{\partial T} = \frac{\partial \bar{U}}{\partial \beta}\frac{\mathrm{d}\beta}{\mathrm{d}T} = -\frac{1}{k_{\mathrm{B}}T^2}\frac{\partial \bar{U}}{\partial \beta} \tag{25}$$

we have

Relationship between heat capacity and energy fluctuations

$$\Delta U^2 = k_{\mathrm{B}}T^2 C \tag{26}$$

It makes physical sense that a system with a large heat capacity should also have large variance of its energy distribution when it is in equilibrium with a heat bath. For, if the heat capacity is large, it means a given influx of energy from the reservoir does not change the system's temperature very much, so the system does not easily become hotter and for that reason cause a heat flow back again.

For many systems at 'high' temperatures (i.e. such that $k_{\mathrm{B}}T$ is large compared to the relevant energy level splittings) one has $\bar{U} \sim CT$ (this is just a rough order of magnitude) and therefore

$$\frac{\Delta U}{\bar{U}} \sim \frac{\sqrt{k_{\mathrm{B}}T^2 C}}{CT} = \sqrt{\frac{k_{\mathrm{B}}}{C}} \sim \frac{1}{\sqrt{N}} \tag{27}$$

where in the last step we asserted $C \sim Nk_{\mathrm{B}}$ which is again a rough statement but applicable to many systems at high temperature. From this we deduce that in thermal equilibrium the standard deviation of the system's energy distribution (often called the energy fluctuation) is about a factor $\sqrt{N}$ smaller than the mean energy.

2 The free energy method

This section can be skipped on first reading. It presents an alternative way to derive the Boltzmann distribution.

In Section 4 of the previous note (STATISTICAL THERMAL PHYSICS: BASIC IDEAS) we derived the Boltzmann distribution as the distribution of probability which maximises entropy under the constraints of fixed average internal energy and fixed total probability. We then considered a pair of (weakly-interacting) systems in thermal contact, and we explored what happens if the apply the Boltzmann distribution to the systems individually (as well as to the composite system overall). We argued that the result makes sense and this justifies the assumption that the constraint associated with mean energy is the right one to describe thermal equilibrium. Arguably, no distribution other than $\{p_i = A_1 \exp(-\beta_1 \epsilon_i),\ p_j = A_2 \exp(-\beta_2 \epsilon_j)\}$ could lead to the maximum entropy of the composite system, but we did not prove this.

We will now present an argument which will underline the conclusion and also provide another way to arrive at it. We begin with the following *thermodynamic* observation: when a system is in thermal equilibrium with a heat bath at a given temperature, the overall entropy of both is maximised when the *free energy* of the system is minimised, where the free energy is given by

$$F = U - TS. \tag{28}$$

The term 'free energy' is a thermodynamic concept with wider application; in the situation of fixed energy levels and number of particles, the free energy is the Helmholtz function.

Using Gibbs's formula for entropy, we have

$$F = \left[\sum_i p_i \epsilon_i\right] + k_{\mathrm{B}} T \left[\sum_i p_i \ln p_i\right]. \tag{29}$$

We would like to find the distribution of probability which will minimise this quantity, with just a single constraint:

$$\sum p_i = 1. \tag{30}$$

Note, there is no constraint on the system energy in this argument, because the system can get any energy it may require from the reservoir. Using a Lagrange multiplier α, the minimum free energy occurs at a stationary value of

$$\sum_i p_i \epsilon_i + k_{\mathrm{B}} T p_i \ln p_i + \alpha p_i \tag{31}$$

By differentiating with respect to p_j at constant $p_{i \neq j}$ and setting the result to zero we find

$$\epsilon_j + k_{\mathrm{B}} T(\ln p_j + 1) + \alpha = 0 \tag{32}$$

and therefore the free energy is minimized when

$$p_j \propto e^{-\epsilon_j / k_{\mathrm{B}} T}. \tag{33}$$

This is the Boltzmann distribution once again. Notice that the parameter β never appeared in this argument. Note also that we explicitly considered thermal equilibrium with a reservoir from the outset, so there was a well-defined temperature and no need to assume anything about the system energy.

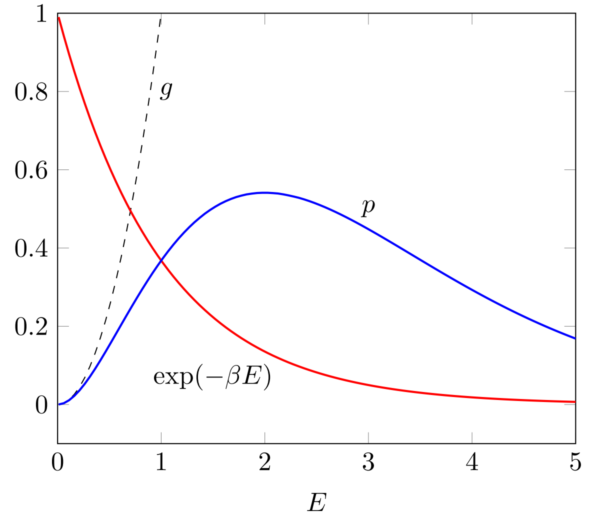

Figure 1: The probability $p(E)$ for the system to visit a given energy level E is given by the product of two factors: g (the number of states of that energy) and $\exp(-\beta E)$ (the Boltzmann factor). The Boltzmann factor decreases with E. For many systems, especially those with many particles, g goes up with E, such that the overall result is a peak in $p(E)$. This peak is narrow for large systems.

3 What is going on in thermal equilibrium

The formulae derived above give most of what we need to know to apply statistical mechanics to a wide range of physical systems. Examples are presented in the next note (SIMPLE APPLICATIONS OF STATISTICAL MECHANICS). It is instructive to see what is going on in all the examples, by considering the distribution of probability over the energy levels of the system.

We can look at the distribution over energy in two ways. The first way is to consider each individual quantum state, labelled by i. In thermal equilibrium the probability for the system to visit that state is

$$p_i = e^{-\beta\epsilon_i}/Z. \tag{34}$$

The second way is more insightful however. What we do is bring in the fact that in practice at any given energy ϵ_i there can be many quantum states with that same energy. The technical word for this is *degeneracy*: we say the energy level has a degeneracy g, where g is equal to the number of quantum states having that energy. (For a system in energy equilibrium at some given energy this degeneracy is precisely the same quantity as W, the number of microstates: $g = W$.)

Now instead of considering individual quantum states labelled by i, we consider groups of quantum states labelled by r, such that the r'th group all have energy ϵ_r. Since probabilities for independent cases add, we have that the probability for the system state to visit the r'th group is

$$p_r = g_r p_i \tag{35}$$

where p_i is the probability of any single member of the group. In other words,

$$p_r = g_r e^{-\beta\epsilon_r}/Z. \tag{36}$$

Thus the probability for the system to have any given energy ϵ_r is given by a product of two factors: the degeneracy factor g_r and the Boltzmann factor $\exp(-\beta\epsilon_r)$. For almost any system of more than a few particles (and also for some single-particle systems), the degeneracy factor is an **increasing** function of energy, whereas the Boltzmann factor is always a decreasing function of energy. So the two factors 'compete': g says the system wants to have lots of energy where there are lots of states, but $\exp(-\beta\epsilon)$ says the system wants to have little energy where the probability per state is high. Overall the system explores all these possibilities and the result is that the probability distribution has a peak somewhere where the increase of g matches the decrease of $\exp(-\beta\epsilon)$. Figure 1 illustrates the idea.

As we consider systems with a larger and larger number of particles N, at some given temperature, the mean energy goes up in proportion to N whereas the standard deviation of the energy distribution goes up in proportion to $\sqrt{N}$ so the relative width of the peak goes down: see (27). In other words the peak gets sharper and sharper, such that the quantities **mean energy** and **most likely energy** agree more and more closely at large N, and also the system state overall is closer and closer to a state of fixed energy. For this reason we often don't trouble to write the bar on quantities such as U and F: it is taken as understood that they refer to the mean value.

4 Exercises

1. Use the expression (14) relating Helmholtz function to partition function to obtain the thermal equilibrium entropy of a harmonic oscillator of frequency ω as a function of $\beta\hbar\omega$.
2. Show that if systems 1 and 2 are weakly interacting but otherwise independent of one another, then the partition function of the joint system is $Z_1 Z_2$. Hence prove by induction that for a system composed of N weakly interacting subsystems all of the same type but distinguishable from one another, $Z = Z_1^N$.
3. Re-derive all the results of Section 1, relying only on your own wits and memory.
4. Consider the simplest possible system: just two quantum states, with energies $\pm\epsilon$.
   (i) Show that the partition function is $Z = 2\cosh\beta\epsilon$.
   (ii) Find the internal energy, first from $\sum_i p_i \epsilon_i$ and then from (5).
   (iii) Show that the heat capacity is $C = k_{\mathrm{B}}(\beta\epsilon/\cosh\beta\epsilon)^2$.
5. Suppose we shift the zero of potential energy, with the result that all the energy levels change by the same amount V. What happens to the partition function, the internal energy and the entropy?
6. System B is like system A except that for every quantum state of A, there are 3 quantum states of B, all of the same energy. How does this difference affect the quantities $Z, \bar{U}, S$?