Derivation of the Boltzmann distribution

A. M. Steane

October 23, 2021

We want to find the values of p_i which maximise

$$S = -k_{\mathrm{B}} \sum_i p_i \ln p_i \tag{1}$$

subject to the constraints

$$\sum_i p_i \epsilon_i = E, \qquad \sum_i p_i = 1. \tag{2}$$

Introduce

$$f \equiv \left(\sum_i \epsilon_i p_i \right) - E \tag{3}$$

$$g \equiv \left(\sum_i p_i \right) - 1. \tag{4}$$

Also introduce Lagrange multipliers $\lambda, , \alpha$ and form

$$y = S + \lambda f + \alpha g. \tag{5}$$

This has a stationary value when

$$\frac{\partial y}{\partial p_j} = 0 \tag{6}$$

where for each j the variables held constant in the partial derivative are all the other $p_{i \neq j}$. Therefore

$$-k_{\mathrm{B}} \ln p_j - k_{\mathrm{B}} p_j \frac{1}{p_j} + \lambda \epsilon_j + \alpha = 0 \tag{7}$$

which gives

$$k_{\mathrm{B}} \ln p_j = \lambda \epsilon_j + (\alpha - k_{\mathrm{B}}) \tag{8}$$

hence

$$p_j = A e^{\lambda \epsilon_j / k_{\mathrm{B}}} \tag{9}$$

where $A = \exp(\alpha / k_{\mathrm{B}} - 1)$. This is the Boltzmann distribution.

There remain two further steps. First we define $\beta = -\lambda / k_{\mathrm{B}}$ and so

---

Boltzmann distribution

$$p_j = A e^{-\beta \epsilon_j} \tag{10}$$

and we define $Z = \sum_i \exp(-\beta \epsilon_i)$. Then the constraint related to α is satisfied when

$$p_i = \frac{e^{-\beta \epsilon_i}}{Z}. \tag{11}$$

Finally, we want to know the value of β. One can show that if two system can exchange energy without a change in their sets of energy levels (hence they are exchanging heat not work) then the entropy of the pair is maximised when they have the same β; this suggests that β is related to temperature. It is not hard to convince oneself that it must be an inverse relationship.

The complete analysis of β is achieved by leaving it as β in the equations obtaining formulae relating U to Z and β, and then F. Eventually we find out that

$$\frac{\partial S}{\partial U} = k_{\mathrm{B}} \beta \tag{12}$$

where in the partial derivative all the energy levels ϵ_i are held constant. But thermodynamic temperature T is equal to $\partial U / \partial S$. Hence we deduce that

$$\beta = \frac{1}{k_{\mathrm{B}} T}. \tag{13}$$

Comments

It is remarkable how the powerful result (10) emerges so quickly from a few simple statements. I think this method of derivation is conceptually one of the most straightforward. In any derivation one has to start by carefully getting one's head around precisely what the symbols mean and what argument is being employed; in the above I think that process is easier than in other approaches.

Note also: there is a spurious argument 'out there' in the physics literature which I will present now in order to refute it. The argument considers a reservoir with a large number of microstates $\Omega(E)$ which is a function of the

---

1

energy E of the reservoir. We suppose this reservoir exchanges energy with our small system whose energy is ϵ. So in order to conserve energy, when the system has energy ϵ the reservoir has energy $E - \epsilon$. The probability of this state of affairs is therefore

$$p(\epsilon) \propto \Omega(E - \epsilon) \times 1 \tag{14}$$

where the 1 signifies the one state consistent with the energy ϵ of the system.

Now, for $\epsilon \ll E$, by using the Taylor expansion we can always write

$$\ln \Omega(E - \epsilon) = \ln \Omega(E) - \epsilon \left. \frac{\mathrm{d} \ln \Omega}{\mathrm{d} E} \right|_E + \dots \tag{15}$$
$$= \ln \Omega(E) - \beta \epsilon + \dots \tag{16}$$

where

$$\beta \equiv \left. \frac{\mathrm{d} \ln \Omega}{\mathrm{d} E} \right|_E . \tag{17}$$

This β is a property of the reservoir. In the limit where the further terms in the Taylor expansion are negligible, we have, then,

$$\Omega(E - \epsilon) = \Omega(E) e^{\beta \epsilon} \tag{18}$$

and therefore

$$p(\epsilon) \propto e^{-\beta \epsilon} . \tag{19}$$

So we have the Boltzmann distribution. Or do we? As I warned you, this argument is spurious as it stands. To be fair, it is not completely spurious, but it is incomplete, and without further statements to justify the approximation of dropping the higher order terms it is almost completely useless.

Here is why.

One can use Taylor series to prove that practically anything is equal to any function you like (to first approximation, that is). For example, here is a "proof" that the log function is equal to the tan function.

Let $y = \ln x$. Introduce $z(x) = \tan^{-1} y(x)$. Now form the Taylor expansion of $z(x)$ about $x = 0$:

$$z(x) = z(0) + x \frac{\mathrm{d} z}{\mathrm{d} x} + \dots \tag{20}$$
$$\simeq A + Bx \tag{21}$$

where $A = z(0)$ and B is the value of dz/dx at $x = 0$. So we have

$$z = \tan^{-1} y = A + Bx \tag{22}$$

hence

$$y = \tan(A + Bx) \tag{23}$$

or in other words

$$\ln x = \tan(A + Bx) \quad ??? \tag{24}$$

The sleight-of-hand here is, of course, that we went from $\simeq$ to $=$ without keeping clear about the approximation involved. The final line should really read

$$\ln x = \tan(A + Bx) + O(x^2) \tag{25}$$

and now everything is alright again. So, in the above (spurious) 'derivation' of the Boltzmann factor what really happened was that we introduced a log function and then made a linear approximation and so ended up with an exponential function. But since we could equally well have arrived at any other function (say by introducing arctan instead of log), it follows that the *whole* of the task of finding out what is the right answer lies in the further work not yet done: the work of showing that the higher order terms vanish in the thermodynamic limit if you pick the log function, but not if you pick other functions. This difficulty did not arise in the Lagrange multiplier method because there the treatment was exact throughout.