

A simple derivation of Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac statistics for systems in a heat bath.

Johan Noldus*

July 15, 2018

Abstract

We derive these laws with a simplicity for high school students and cocky beach girls.

1 Maxwell-Boltzmann.

Consider a system of N particles in contact with a heat reservoir; the system can be in a finite number of states $|i\rangle$ each characterized by an energy ϵ_i and other numbers α_i . The question is, given a temperature T of the heat bath, what is the probability that the system is found in $|i\rangle$? One might argue that the question is ill posed given that we always find the system in a certain state and that after that measurement, it cannot gear up sufficiently fast to be in the pre-measurement state again. This argument cannot be sidestepped by assuming that the heat tank is infinitely large, only its temperature would determine response times and even then an extremely fast post measurement intervention would grant the tank with an intelligence it does not possess. Therefore, given that we have to allow for the system to change from $|i\rangle$ to $|j\rangle$ with a probability $p(i \rightarrow j)(t)$ per unit time and set that the only characteristic known of the heat bath is its temperature, we must find that $p(i \rightarrow j)(t)$ is only dependent upon the characteristics of the system and T and nothing else. These probabilities have to be determined dynamically; suppose initial probability $p_i(t)$ is given. The latter satisfies

$$\dot{p}_i(t) = \sum_j (p_j(t)p(j \rightarrow i)(t) - p_i(t)p(i \rightarrow j)(t))$$

where t is time. One notices that

$$\sum_i \dot{p}_i(t) = \sum_{i,j} (p_j(t)p(j \rightarrow i)(t) - p_i(t)p(i \rightarrow j)(t)) = 0$$

*email: johan.noldus@gmail.com, Relativity group, departement of mathematical analysis, University of Gent, Belgium.

so that total probability is conserved. Given that we have two unknowns $p_i, p(i \rightarrow j)$, it is clear that we need a secondary law. The obvious, and strongest, candidate is that

$$p(i \rightarrow j)p(j \rightarrow k) = \kappa p(i \rightarrow k)$$

must be j independent and $\kappa(i, k)$ independent as a matter of ‘‘homology’’ condition. It signifies that, in a way, the transition from i to k can happen in multiple stages where only the number of intermediate, unknown, stages matters and not the details thereof. The stronger form sets κ equal to one, which would mean that a system has its own radiative temperature and cannot be heated up, by means of a reservoir, to a stable state of higher temperature. This law can be derived in another way by noticing that

$$p(i \rightarrow j) = p_{i^*} p_j$$

where p_i has some functional form in terms of ϵ_i and α_i and p_{i^*} is the fictional probability of destroying a state $|i\rangle$. This means that the probability of transition is given by the probability of destruction of a state $|i\rangle$ followed by the birth of $|j\rangle$ where both these happenings are independent of one and another. Obviously, this is achieved by putting ϵ_i, α_i to $-\epsilon_i, -\alpha_i$ and one remarks that

$$\dot{p}_i(t) = \left(\sum_j p_j(t) p_{j^*}(t) - p_{i^*}(t) \right) p_i(t).$$

Now, an energy ϵ_i reflects a certain wavelength λ_i with $\epsilon_i = \frac{hc}{\lambda_i}$ and c the speed of light, \hbar the quantum constant. $\tau_i = \frac{\lambda_i}{c} = \frac{\hbar}{\epsilon_i}$ is a timescale associated to observation of that energy. A suitable defining characteristic is therefore

$$0 = p_i\left(\frac{\hbar}{\epsilon_i}\right) - p_i(0) = \int_0^{\frac{\hbar}{\epsilon_i}} \dot{p}_i(t) dt = \int_0^{\frac{\hbar}{\epsilon_i}} \left(\sum_j p_j(t) p_{j^*}(t) - p_{i^*}(t) \right) p_i(t) dt.$$

For example, $p_i(0) = 1$ and $p_j(0) = 0$ with $i \neq j$ satisfies this criterium although $p_i(\tau) = 1$ for all $\tau > 0$ and therefore this timescale is rather ambiguous. It is utterly clear, given that the system can only ‘‘sing’’ the modes ϵ_i that this requirement signifies that it is stable on the associated timescale. This definition of temporary temperature $T_{[0, \tau_i]}$ associated to a time interval of measurement, for example by putting your finger on a heating plate, is now open for discussion and weakening. Now, we come back to the homology law which would imply that

$$p_{j^*} p_j = \kappa$$

for all j and therefore

$$\dot{p}_i(t) = N \kappa p_i(t) - \kappa.$$

This leads for a constant κ to solutions of the kind

$$p_i(t) = \frac{1}{N} (1 - a_i(\epsilon_j, T)) e^{N \kappa t}$$

which leads to

$$\sum_i a_i(\epsilon_j, T) = 0$$

and restricted measurement times as not to get negative probabilities. In general, it constitutes a balancing around the uniform distribution which

is in conflict with $p_{j^*}p_j = \kappa$ as one can easily check. This suggests one to replace

$$p(i \rightarrow j)p(j \rightarrow k) = \kappa p(i \rightarrow k)$$

with

$$\sum_j p(i \rightarrow j)p(j \rightarrow k) = \kappa p(i \rightarrow k)$$

which is a principle of ignorance indicating that one does not know the intermediate state as a matter of principle. This leads to

$$\sum_j p_{j^*}p_j = \kappa$$

and therefore

$$\dot{p}_i(t) = (\kappa - p_{i^*}(t))p_i(t).$$

This at least incorporates the case $p_i = 1, p_j = 0$ given the functionality $p_k = \delta(\epsilon_k - \epsilon_i)$.

Traditional Maxwell-Boltzmann assumes the probability distribution is time independent which, in principle, is very wrong but good enough for sufficiently high temperatures given that “adiabatization” occurs sufficiently fast. Given that last assumption, our time independent distribution must satisfy the property that it factorizes over different “independent subsystems”. That is

$$p_{1 \cup 2}(\epsilon_i^1 + \epsilon_j^2) = \sum_{(k,l): \epsilon_k^1 + \epsilon_l^2 = \epsilon_i^1 + \epsilon_j^2} p_1(\epsilon_k^1)p_2(\epsilon_l^2)$$

and $p(\epsilon_i)$ where ϵ_i is a time independent energy. This implies

$$p(\epsilon_i) = \frac{e^{-\beta\epsilon_i}}{\sum_j e^{-\beta\epsilon_j}}.$$

Here, $\beta = \frac{1}{k_B T}$ where k_B is Boltzmann constant and T the temperature in kelvin. This is *not* a solution to our above system and the physical significance of temperature in terms of more mundane units is that

$$\langle \epsilon \rangle = \sum_i \epsilon_i p_i = k_B T^2 \frac{d}{dT} \sum_j e^{-\beta\epsilon_j}.$$

The average energy divided by $k_B T^2$ is therefore the slope of the state sum when varying the temperature. The latter is a kind of number associated to the system which cannot be measured. The separability condition is often assumed to be correct although it excludes hidden correlations between both subsystems by means of interaction through the heat bath. Such equilibrium will never settle as is most easily seen which validates our previous assumptions. In both attempts to define temperature, a characteristic time or length scale emerged, the second one which is associated to infinite observation times and heat supplies.

2 Ideal gas law.

Finally, let me present a most easy derivation of the ideal gas law. Every system contains order and disorder where the latter is defined in the socialist sense meaning that this energy is useless for labour. Therefore, any

robot(a) doing work must be sufficiently hot to do so. Woman are cold so they are passive during sex whereas hot men can only arise Hercules. We assume that the energy of warmth per unit rabota is proportional to its temperature in Kelvin by means of Boltzmann constant; henceforth, we have formula of Vladimirovna Romanova that

$$E_{\text{internal energy minus heat}} := Nk_B T$$

where N is the number of workers. Since labour in Russia is expressed in terms of machine pressure and volume (excluding intelligent activities) of steam engines, we we have Lenin formula given by

$$E_{\text{Lenin}} = pV.$$

Now, worker must be hot in order to do something (useless) so that after work process heat always remains; henceforth

$$0 \leq E_{\text{internal energy minus heat}} - E_{\text{Lenin}}.$$

Vladimirovna's paradise is, justifiably for Russians, given by the limit of this formula, excluding worker intelligence and increasing Romanov safety, leading to

$$pV = Nk_b T$$

which is ideal gas law.

As far as general laws go, we now introduce the notion of Shannon entropy which is a bit of a cheat. We start from a free system in equilibrium; experience dictates that to a very good approximation

$$E = TS - pV$$

where E is the internal energy and S the so called entropy, a function Maxwell and Von Neumann were looking after. In a crude way, this formula makes sense, if pressure increases, system loses energy due to labour, similarly if volume increases. If temperature goes up, then internal energy goes up linearly, so what is S ? Lets calculate it from Maxwell Boltzmann and ideal gas law:

$$E = \sum_i E_i P_i = N \sum_i \epsilon_i p_i = TNs - \frac{N}{\beta}$$

where the small latin letters pertain to a single particle and the big ones to the entire gas (here we use the factorization property of the distribution). It is clear that $S = Ns$ for a free gas given that V, S must be extensive quantities, meaning that they add up when two systems are joined in one system keeping all other parameters fixed. More generally $S_{1 \cup 2} = S_1 + S_2$ for noninteracting systems one and two. Henceforth

$$s = k_B + \frac{1}{T} \sum_i \epsilon_i p_i$$

where $\epsilon_i = -k_B T (\ln(p_i) + \ln(\sum_j e^{-\beta \epsilon_j}))$. Therefore

$$S = -k_B \sum_i p_i \ln(p_i) + k_B (1 - \ln(\sum_j e^{-\beta \epsilon_j}))$$

and for a *free* particle at temperature T , $\sum_j e^{-\beta \epsilon_j}$ must be a dimensionless positive constant, which can be reset, by a translation on the energy scale

to e . Only entropy has this freedom given that all other intrinsic T, p and extrinsic V parameters are fixed. Henceforth

$$s = -k_B \sum_i p_i \ln(p_i)$$

which is the correct definition of Shannon and Von Neumann and satisfies invariance under energy translations, in either it does not directly depend upon them. Having said this, it must be clear that s only works at best in Vladimirovna's paradise and simple energy relations of this kind.

The ultra cocky beach girls might now utter that the property $S_{1 \cup 2} = S_1 + S_2$ defines the entropy function in case of *distinguishable* particles. Here, every (i, j) determines an energy $\epsilon_i^1 + \epsilon_j^2$ and other pairs with identical sums are treated separately. Given that products split into sums, we need logarithms;

$$S_{1 \cup 2} := \sum_{i,j} \ln(p_i^1 p_j^2) := M S_1 + N S_2$$

where M, N is the number of states in 2, respectively 1. To correct for this feature, we must settle for

$$S = -\kappa \sum_i p_i \ln(p_i).$$

Positivity of S determines that $\kappa > 0$ because the natural logarithm is negative on the interval $[0, 1]$. The reader should investigate the function $x \rightarrow -x \ln(x)$ a bit further. It has two zeroes 0, 1 and a maximum at $\ln(x) = -1$ or $x = \frac{1}{e}$. Henceforth it is concave, meaning that the functional value of a point between two points is above the value of the straight line connecting both points in the plane.

Reversly, we might infer now the ideal gas law from the first law of thermodynamics $E = TS - pV$ as well as the factorization property and positivity of the entropy as well as the probability distribution itself. In thermophysics, we have the assumption of no intelligence, meaning that adding two systems together is a subadditive process regarding entropy, especially when they remain in some contact with one and another. That is,

$$S_{1 \oplus 2} \geq S_1 + S_2$$

where the \oplus signifies that they are not necessarily disjoint and still in contact with one and another. All this is rather vague and not applicable in the world of intelligent creatures.

3 Einstein and Dirac distributions.

Maxwell Boltzmann is not always applicable in nature, there are circumstances where individual particles lose their identity in a group identity of identical particles. There are two ways of doing so, the Bose way or the Fermi way. To explain this mathematically is very simple and requires a bit of quantum theory: the "states" of a single system form a complex vector space. This means that if x, y represent states of S , then $ax + by$ also does in a way, where $a, b \in \mathbb{C}$ are complex numbers. This is the principle of superposition which is the pinnacle of quantum theory. Given two states x, y , then it is natural to posit that one has an action

of the exchange operator $R(x, y) = (y, x)$ such that any identity from x, y satisfies

$$R(I(x, y)) = aI(R(x, y)) = I(y, x)$$

where a is a complex number so that $R(I(x, y))$ and $I(R(x, y))$ are the same upto a multiple. Since $R^2 = 1$, we have that $a = \pm 1$ and therefore $RI \mp IR = 0$; the first option is called Bose-Einstein, the second Fermi-Dirac. The reader easily calculates that $I_+(x, y) = c((x, y) + (y, x))$ and $I_-(x, y) = c((x, y) - (y, x))$ where c is a complex number. The reader is invited to generalize this to n bosonic and fermionic particles. So, in case for M Bosons, the energy state,

$$\epsilon_M := \sum_{k=1}^M \epsilon_{i_k}$$

arises

$$\frac{M!}{\prod_{j=1}^N (N_j)!}$$

times where $\sum_{j=1}^N N_j = M$ and N_j is the faculty of the number of times ϵ_j appears in the sum ϵ_M with the agreement that $0! = 1$. The probability distribution is henceforth given by

$$p(\epsilon_M) := \frac{\frac{M!}{\prod_{j=1}^N (N_j)!} e^{-\beta \sum_{k=1}^M N_k \epsilon_k}}{\left(\sum_{j=1}^N e^{-\beta \epsilon_j}\right)^M}$$

assuming that every energy level is uniquely determined by (N_j) . For big N_j ,

$$\frac{e^{-\beta N_j \epsilon_j}}{(N_j)!} \sim A \left(\frac{e^{-\beta \epsilon_j}}{N_j} \right)^{N_j}$$

by means of Stirling formula. The reader is invited to work out this formula further by taking the natural logarithm and importing entropy function.