

2/12/10

Dr. Laurendeau,

Since my letter of 10/27/09, I have given more thought to your rationale in section 3.4 of *Statistical Thermodynamics*. You're surely aware that in the context of your analysis, the most probable distribution is given by

$$W_{mp} = \frac{N!}{(N/M)!^M}, \quad (1)$$

as illustrated by your example showing that $W_{mp} = W_d\{3,3\} = 20$ (p. 36). However, if you attempted to compare W and W_{mp} directly you may have encountered the same computational difficulty as I did upon applying the short form of the Stirling approximation. You attempt to circumvent this problem by observing that W_{mp} is bounded by W and W_{avg} . (I've modified the notation here due to the limitations of my word processor.) Then you reason that if W and W_{avg} converge, W and W_{mp} must also. However, it is clear from the unnumbered equations in the middle of page 36 that W and W_{avg} do *not* converge since the ratio $W_{avg}/W = 1/(N+1)$, which approaches zero with increasing N . (The fact that the ratio of the *logarithms* approaches 1 is clearly irrelevant. Logarithms are usually added/subtracted, as illustrated by the first paragraph on page 41.)

While the above error undercuts the rationale for your derivation in section 3.6, Eq. (3.20) can be justified instead by adopting Boltzmann's assertion that an isolated system will spontaneously evolve to its most probable macrostate W_{mp} . However, a return to the beginning of chapter 3 reveals a deeper problem, which is framed at the end of the first paragraph, when you declare that "by invoking the assumption of independent particles, our upcoming statistical analysis can be based rather straightforwardly on probability theory describing independent events". You then proceed to present the "two basic postulates of statistical thermodynamics" (ergodicity and equiprobability), followed by the two system constraints (conservation of mass and energy). On page 34, you state that "if each system quantum state is equally likely, then every microstate must also be equally likely." Of course, the converse is also true.

However, one can easily see that the quantum energy states can only be equiprobable under the very special circumstance where only one energy state is available with an energy ϵ equal to the average particle energy E/N . If we postulate another available energy state either greater than or less than the average energy, then the equiprobability assumption implies the existence of a microstate where all of the particles are in that energy state. However, such a microstate would clearly violate energy conservation and therefore would not be possible. But eliminating that microstate would also eliminate the corresponding occasions of particles being in that energy state, reducing the overall probability of a particle being in that energy state.

A similar situation is illustrated by your Example 3.1. Since you have correctly eliminated all macrostates that violate energy conservation, it is clear from this example

that under the assumption that the allowable microstates are equally probable, the probabilities of being in energy levels ε_0 , ε_1 and ε_3 are .3, .4, and .3 respectively. Including the degenerate states gives probabilities of .3, .2, .2, .1, .1 and .1, reading from bottom to top. Of course, if the energy states are not equally probable, then the microstates cannot be either and the combinatorial microstate counts can no longer represent macrostate probabilities.

Contrary to the treatment of Example 3.1, the example of section 3.4 ignores the energy conservation constraint. Why is this? It would seem logically consistent to apply the constraint to this case as well. If we apply the conventions of Example 3.1 and assume $M = 2$ energy levels of $\varepsilon_0 = 0$ and $\varepsilon_1 = 1$, we will have 7 possible scenarios with system energies 0 through 7. Each of these corresponds with one of the 7 macrostates, so that if we select, for instance, $E = 4$, the corresponding macrostate will be $\{2, 4\}$. The other macrostates will be eliminated, as in Example 3.1, since they violate energy conservation. Therefore, the only value of E for which equiprobability of energy states holds is $E = 3$, corresponding to the macrostate $\{3, 3\}$. Furthermore, if we follow the procedure dictated by your postulates and constraints, the formulas for W_m , and W given in section 3.4 are meaningless, just as they are for Example 3.1. Instead of $W_m = 7$, we get $W_m = 1$, and instead of $W = 64$, we get various values between 1 and 20, depending on the system energy E . However, we can now be sure that $W_{mp} = W$, since there is only one allowable macrostate for each selected value of E .

Statistical mechanics has been assembled on the edifice of combinatorial analysis, as described in chapter 2. The core assumption of combinatorial analysis is that of equiprobability, meaning that the individual particle states are treated as “*mutually exclusive*, equally likely points in *sample space*” (p. 7). However, even the most cursory *a priori* analysis shows that points in energy (or velocity, momentum, etc.) space cannot be equally probable, which means that the related microstates are also not equally probable and macrostate “probabilities” derived from combinatorial analysis are meaningless. This is of course not a problem if one is considering the *spatial* distribution of molecules in an ideal gas or the distribution of particles within a single energy level, since these distributions are not constrained by energy conservation. However, the primary goal of chapter 3 is to derive a distribution specifying different probabilities for each energy level, which inherently contradicts the core premise of equiprobability.

The root cause of this contradiction lies with the “assumption of independent particles”. This assumption is clearly false if the system is constrained by conservation of energy since the increase in the energy of any particular particle must necessarily be offset by an equal decrease in energy of the rest of the system. In fact, this applies to any pair of interacting particles. For instance, if we take the form of the interaction to be a classical collision between two particles, the sum of the kinetic energies of the particles before the collision must be the same as the sum after the collision, which means that the energy of one particle is strictly conditioned on the energy of the other. More formally, if we take proposition A to be “particle 1 has energy ε_0 ” and proposition B to be “particle 2 has energy ε_1 ”, then according to the logic of probability theory

$$P(A, B) = P(A)P(B|A). \quad (2)$$

Independence requires that

$$P(A, B) = P(A)P(B), \quad (3)$$

which is only true if

$$P(B|A) = P(B). \quad (4)$$

But for a two-particle system with total energy $\varepsilon_T = \varepsilon_0 + \varepsilon_1$, we know that if A is true, then B must be true also, so that $P(B|A) = 1$. However, the assumption of equiprobability dictates that the unconditional probability $P(B) = 1/2$ (assuming that ε_0 and ε_1 are the only possible energy states), rendering the equality of eq. (4) false. So the fundamental premise of statistical thermodynamics is in direct contradiction to the principle of conservation of energy.

While the Boltzmann method for deriving the energy distribution gets the desired answer, it would be nice if we could avoid the above contradiction. One approach is to recognize that Boltzmann's "relative probability" is just the multinomial coefficient of the multinomial distribution, as discussed in your Chapter 2. The complete expression for the probability of a classical macrostate is given by $P_m = W_B P_\mu$ where W_B is equivalent to the W_{MB} from your Eq. (4.7) with no degeneracy. W_B and P_μ are then

$$W_B = \frac{N!}{\prod_{i=1}^M N_i!} \quad (5)$$

and

$$P_\mu = \prod_{i=1}^M P_i^{N_i}. \quad (6)$$

P_μ corresponds to the *microstate* probability, which Boltzmann assumes to be constant with variations in $\{N_i\}$. It is clear from eq. (6) that that this can only be true if $P_i = 1/M$ for all i . However, we know this cannot be the case, and are in fact trying to derive a contrary result. Conservation of energy must hold for all microstates; maximizing the dispersion of energy within a microstate corresponds to minimizing P_μ subject to the energy conservation constraint. (This is isomorphic to minimizing the denominator of W_B in the Boltzmann derivation.)

Taking the logarithm of eq. (6) gives $\ln P_\mu = \sum N_i \ln P_i = N \sum P_i \ln P_i$, which when fed into the Lagrange method yields

$$P_i = e^\alpha e^{\beta \varepsilon_i} . \quad (7)$$

Setting $\beta = -1/kT$ and $e^\alpha = 1/Z$ gives

$$P_i = \frac{e^{-\varepsilon_i / kT}}{Z} . \quad (8)$$

The values of $\{P_i\}$ are the probabilities which correspond the maximum allowable dispersion of energy. This derivation avoids the complications of combinatorial mathematics and therefore has no need of the Stirling approximation.

From the above we see that for the macrostate probability $P_m = W_B P_\mu$ to be *stationary*, W_B must be maximized while P_μ is minimized. This suggests a *minimax* approach with a saddle point solution which satisfies the equation $dP_m = W_B dP_\mu + P_\mu dW_B = 0$. Taking P_μ as constant corresponds to the Boltzmann assumption of equal microstate probabilities, so that $dP_\mu = 0$. The above equation reduces to $P_\mu dW_B = 0$, which is then solved subject to the energy conservation constraint to obtain the optimal energy distribution. This distribution can be described in terms of either $\{N_i\}$ or $\{P_i\}$ by assuming that the probabilities can be inferred from the optimal frequency distribution of the particles so that $P_i = N_i/N$ for all i . The alternative P_μ approach is to assume W_B is constant, so that $dW_B = 0$. This yields $W_B dP_\mu = 0$ to be solved subject to the energy conservation constraint. While both approaches yield the same solution, minimizing P_μ allows for a more logically consistent rationale.

The derivation from P_μ is reminiscent of Einstein's approach in his famous photoelectric paper [Concerning an Heuristic Point of View Toward the Emission and Transformation of Light, 1905]. In section 5 he asks "How large is the probability that at a randomly chosen instant of time all n movable points in the given volume v_0 will be found by chance in the volume v ?" His answer is $W = (v/v_0)^n$, which can be relabeled with notation similar to yours as $W = (V/V_0)^N$. Assuming that the particles are distributed uniformly over the volume V_0 regardless of energy, each particle will on average occupy a volume $V_p = V_0/N$. Therefore, the particles with energy ε_i will occupy a volume $V_i = V_p N_i = V_0 N_i / N$. The probability of a particle being in energy state i is then $P_i = N_i/N = V_i/V_0$, so that the probability of a given distribution of energy over M energy states is

$$P_\mu = \prod_{i=1}^M P_i^{N_i} .$$

The above approaches all rely on the Lagrange method, which requires a rationale for maximization/minimization. Boltzmann's "laws of probability" are insufficient for this purpose, since the extent of our knowledge regarding the state of the system can have no physical effect on the motion of the particles. Your commendable attempt to avoid this

difficulty by pursuing a different rationale also falls short, so we are left with no credible justification for applying the Lagrange method.

Once again, Einstein hints (perhaps inadvertently, in footnote 6 referenced from the above section 5) at how we might derive the energy distribution while avoiding the need for the Lagrange method as well as the logical difficulties presented by the combinatorial approach. This approach involves the following steps for deriving the distribution of the energy states over the volume of an ideal gas.

$$pV = NkT \quad \text{for an ideal gas at equilibrium.} \quad (\text{S1})$$

$$dU = -p dV \quad \text{for an adiabatic system with internal energy } U. \quad (\text{S2})$$

$$dU = -NkT dV/V \quad \text{from combining (S1) and (S2).} \quad (\text{S3})$$

$$dV/V = -dU/NkT \quad \text{rearranged.} \quad (\text{S4})$$

$$\ln |V| + C = -U/NkT \quad \text{integrated.} \quad (\text{S5})$$

$$V = \exp[-U/NkT + C] = A \exp[-U/NkT] \quad (\text{S6})$$

$$U_i = N_i \varepsilon_i \quad \text{where } \varepsilon_i \text{ is the energy of the } i\text{th energy state and } U = \sum U_i \quad (\text{S7})$$

$$V_i = A \exp[-\varepsilon_i/kT] \quad \text{where } V_i \text{ is the volume containing the } N_i \text{ particles} \quad (\text{S8})$$

Assuming that at equilibrium the particles are distributed uniformly over the volume $V = \sum V_i$ results in

$$P_i = \frac{N_i}{N} = \frac{V_i}{V} = \frac{Ae^{-\varepsilon_i/kT}}{V}, \quad (\text{9})$$

which normalizes to eq. (8) by setting $Z = V/A$. The assumption implicit in classical gas theory is that the particles are uniformly distributed over physical space, which is also assumed for the derivations involving P_μ . The *a priori* assumption of equiprobability over physical space is more plausible than Boltzmann's *a priori* assumption of equiprobability over energy states and avoids the associated logical contradiction.

All of the above derivations assume that a stationary equilibrium is spontaneously approached, but all are devoid of any physically causal explanation of why this occurs. As Einstein points out (in the last sentence of the aforementioned section 5) "It is noteworthy that in the derivation of this equation [for entropy]... no assumption had to be made as to a law of motion of the molecules." This observation applies to the above derivations as well. However, the physical cause of the tendency of systems to evolve toward equilibrium cannot be established without specifying the underlying dynamical mechanism(s) which lead to this behavior.

As you can see, I have found your book quite stimulating. My comments are intended to be constructive and I would appreciate any response you may have.

Best regards,

Bill Dreiss