

## Degeneracy, Average Occupancy and Microstates in Statistical Mechanics

Francesco R. Ruggeri Hanwell, N.B. May 15, 2023

An interesting result arises in information theory, namely that if one considers a microstate with large  $N$  trials, each representing a state  $i$  with probability  $p(i)$ , then even though this is only one of many possible microstates (each being a permutation of each other) it carries information about all of the permutations based on the average occupancy number. In particular, the probability of a microstate may be written as:  $\prod_i p(e_i) [p(e_i)^{N p(e_i)}]$  for large  $N$ . This holds for any microstate and so all have the same probability. Given that all of the microstates are permutations in the large  $N$  case, this probability is equivalent to  $1/\{\text{number of permutations}\}$  where  $\text{Number of permutations} = N! / \{\prod_i n(e_i)!\}$ .

We point out the above example, because it seems there are at least two main strategies in statistical mechanics. One is to begin with a set of average occupancy numbers  $n(e_i/T)$  whose specific form is unknown. The idea then seems to be to seek degeneracy in the system, possibly among the arrangement of the  $n(e_i)$ 's or degeneracy within single levels  $e_i$  and to maximize the number of degenerate arrangements in order to account for as little information as possible. This approach may be used to calculate the Maxwell-Boltzmann, Fermi-Dirac and Bose-Einstein distributions. In the case of the latter two, degeneracy in the single energy levels  $e_i$  is key. One then maximizes the  $\ln$  of the degenerate arrangements subject to constraints such as  $\sum_i e_i n(e_i)$  and  $\sum_i n(e_i) = N$ . The degeneracies themselves are associated with momentum degeneracies yielding the same energy.

A completely different approach, it seems, is to consider  $\exp(-(E-uN)/T)$  as representing the probability of a state with energy  $E$  and  $N$  particles and to use either the grand canonical ensemble i.e. sum the various  $\exp(-(E-uN)/T)$ 's. These may then be broken into single particle sums using  $E(N) = \sum_i e_i n(e_i)$  for  $n(e_i)$  variable or  $\sum_i e_i$  for  $N$  fixed. Even though momentum degeneracy still exists, it is not part of the  $n$  sum of  $\exp(-(E-uN)/T)$  factors which gives a single energy level probability distribution with no maximization necessary. Thus we point out that  $\exp(-(E-uN)/T)$  for microstates seems to incorporate information about degeneracies without them needing to be explicitly added. On the other hand, if one starts with average  $n(e_i)$  (without knowing their form), one must explicitly consider degeneracies in order to maximize so that there is no unnecessary information in the system. Both approaches should yield the same result. We consider the Bose-Einstein example of a photon gas as an example.

### Microcanonical States

In general, one deals with a system of  $N$  microcanonical particles with a total energy  $E$  microcanonical. One may consider all different ways to distribute this energy (called microstates), but mathematically it is difficult in large  $N$  cases to impose the constraint of a fixed  $E$ . Thus this constraint is relaxed and one considers all total energy scenarios with the weight  $\exp(-E/T)$ . This is said to lead to a sharp distribution about  $E_{\text{average}} = E_{\text{microcanonical}}$  for an appropriate  $T$ . To be more general, one may also consider the factor  $\exp(-uN/T)$  to introduce

systems with more than  $N_{\text{microcanonical}}$  numbers of particles. This again should be a sharp distribution for  $u, T$  with the average  $N$  being  $N_{\text{microcanonical}}$ .

A main idea is that:

$$E = \sum_i e_i n(e_i) \quad ((1))$$

Thus a sum over  $E$  states may be replaced with a product of single particle sums i.e.

$$\sum \exp(-E/T) = \prod_{j=1}^N \left\{ \sum_i \exp(-e_i/T) \right\} \quad ((2))$$

((2)) holds for fixed  $N$ . One may see that it includes all arrangements of  $e_i$ 's from the  $N$  different particles i.e. the permutations. In terms of momentum degeneracies associated with  $e_i = p^2/2m$ , these may be introduced by replacing:

$$\sum_i \text{ with } \rightarrow \text{ constant } \int d^3p \quad ((3))$$

In ((2)) the  $\{ \}$  represents a sum over possible energies for a single particle. We note that the single particle probability  $\exp(-e_i/T)$  ensures that any microstate of  $e_i$ 's adding to the same total energy has the same probability. We have already noted that the various permutations of the  $e_i$ 's are included in this approach.

An important question which arises is: Why should one use  $\exp(-E/T)$  as the probability for the total state? We note that  $\exp(-E/T)$  follows from considering a system with many states which accounts for the minimum information necessary. In (1) it is argued that the number of states less than or equal to  $E_{\text{total}}$  is proportional to  $E_{\text{total}}^f$  where  $f$  is the number of degrees of freedom in the system i.e. proportional to the number of particles. If one uses a power law distribution  $(E_{\text{total}} - E)^f$ , then

$$\text{States between } (E - \delta \text{ and } E + \delta) \text{ proportional to } \delta (E_{\text{total}} - E)^f \quad ((4))$$

In the high  $N$  limit, ((4)) becomes a Maxwell-Boltzmann weight for  $E$  i.e.  $\exp(-E/T)$ .

As noted,  $E$  may be written as a product as in ((2)) with degeneracies of  $p$  included in ((3)).

One may note that each degeneracy constant  $d^3p$  is associated with a different single particle energy level. The sum  $\{ \}$  in ((2)) only considers one particle in each product term.

### Microcanonical States with $N$ flexible

Let us now consider the  $\exp(-E/T)$  weight applied to a system of photons with a single frequency  $\hbar \omega$ . In such a case, one cannot use:

$$\left\{ \sum_i \exp(-e_i/T) \right\} \text{ where } i \text{ is over different single particle energy levels.}$$

Here there is only one energy level  $\omega$ , but one may have different numbers of photons in this level. In such a case:

Sum over  $\exp(-(E/T - uN/T))$  becomes:

Sum over  $n \exp\{-(\hbar\omega - u) n/T\}$  where  $n$  is the number of photons in a “microstate” ((5))

A main idea we wish to point out is that when summing over  $e_i$  states, a single energy level may have degeneracy: constant  $d^3p$ . A state with  $n$  photons should also have momentum degeneracy (i.e. momentum vectors pointing in various directions), but the sum over  $n$  does not include these. The sum over  $n$  is simply a discrete sum which ignores the degeneracy. Thus:

Ln Partition function  $Z = \ln\{ \text{Sum over } n=0, \text{ infinite } \exp\{-(\hbar\omega - u) n/T\} \}$  ((5)) is a geometric series

$\text{approx} = \ln\{ 1 / [1 - \exp\{-(\hbar\omega - u) / T\}] \}$

Average occupancy =  $n(\omega) = -Td/d(\hbar\omega) Z = 1 / \{ \exp\{(\hbar\omega - u) / T\} - 1 \}$  ((6))

((6)) is the Bose-Einstein distribution.

Thus ((6)) is obtained without any use of degeneracies in the direction of momentum. If one had a photon gas with different  $\omega$ , then one could consider a integral over constant  $\int d^3p$  where  $|pc| = \hbar\omega$ .

One does not need to use the photon example with a single energy  $\hbar\omega$  to obtain a Bose-Einstein distribution using the grand canonical approach. In (2) (page 186)  $n$  sums again lead to geometric series which create the basic form of the Bose-Einstein distribution.

## Average Occupancies and Degeneracies

A completely different approach to solving for average occupancy is to:

((A)) Assume average occupancy  $n(e_i)$  value for each  $e_i$  single particle level. These are completely unknown and not necessarily linked to  $\exp(-E/T)$

((B)) Calculate degeneracies in the system. One has to be careful to only include a particular type of degeneracy once. For example, in the Maxwell-Boltzmann case, one must consider that all  $n(e_i)$  particles in a level  $n(e_i)$  are identical. This has the overall appearance of considering all permutations of a set  $\{n(e_i)\}$ . In the Bose-Einstein case, one considers degenerate arrangements in each single  $e_i$  state with degeneracy  $g_i$  and this calculation already contains the identical nature of the  $n(e_i)$  particles. Thus one does not introduce  $N! / \text{Product } n(e_i)!$  a second time.

((C)) Take  $\ln$  of the number of permutations/degeneracies and maximize (minimize information) subject to constraints such as:  $\text{Sum over } i e_i n(e_i) = E_{\text{ave}}$  and  $\text{Sum over } n(e_i) = N_{\text{ave}}$

In this approach, degeneracies are key even though they did not appear at all in the procedure of obtaining ((6)). This, we argue, is curious and we consider later. First we give two examples.

The Maxwell-Boltzmann approach begins with  $n(e_i)$  single particles with energy  $e_i$  and  $N$  total particles. Each  $e_i$  is associated with a degeneracy  $g_i$  from momentum, but the overall momentum degeneracy is taken as a product of  $g_i$  (just as in the single particle situation in the above section). This, however, is not a correct counting of degeneracies for identical particles. For example, consider two particles in two possible degenerate states. One may have both in the first state, both in the second and one in each. This leads to 3 possibilities, but  $g_i=2$  and  $g_i$  to the power  $n=2$  is 4 i.e there is overcounting. Thus the MB case is an approximation. Furthermore, two types of degeneracies are considered. First, there are the permutations of the  $n(e_i)$  particles and secondly the momentum degeneracies. Using:

$$\ln\{ N! / \text{Product over } i \, n(e_i)! \} + \text{product over } i \, g^{n(e_i)} \quad ((7))$$

as the number of degeneracies and maximizing with respect to  $n(e_i)$  using the constraint  $\sum_i n(e_i) = N$  yields:

$$P(e_i) = C g_i \exp(-e_i/T) \quad ((8))$$

Next, consider taking single particle degeneracies more seriously. Again assume  $n(e_i)$  particles in each  $e_i$  single level on average with  $g_i$  being the degeneracy in this level. One may consider permutations of  $n(e_i) + g(e_i) - 1$  objects with  $g_i - 1$  being partitions. For example, for  $g_i=2$ , only 1 partition is needed. Then the degenerate permutations within each level are:

$$(n(e_i) + g_i - 1)! / \{ n(e_i)! (g_i - 1)! \} \quad ((9))$$

Taking  $\ln$  of ((9)) and maximizing with respect to  $n(e_i)$  with the constraint  $\sum_i (e_i - u) n(e_i) = 0$  yields the Bose-Einstein distribution:

$$n(e_i) = \text{average occupancy of } i = g_i / \{ \exp( (e_i - u)/T ) - 1 \} \quad ((10))$$

One may note that one does not consider degeneracies  $g_i$  and permutations of the  $n(e_i)$ 's among themselves as in the MB case, but only ((9)). It may be noted that the information indicating that the  $n(e_i)$  particles are identical is already included in ((9)) in the term  $n(e_i)!$  in the denominator. Thus this information is being removed and cannot be removed a second time by using  $N! / \text{Product } n(e_i)!$  as well. In the MB case, one has to indicate both  $g_i$  degeneracy and the fact that the  $n(e_i)$  particles are identical through  $n(e_i)!$ . Thus the factorial approach really tries to remove information which does exist i.e. the identities of the particles.

We ask the question: Why are degeneracies excluded in the sum over  $n$  in the  $\exp(-E/T)$  approach to photons while the factorial approach forces one to manually calculate the degeneracy expression in order to find  $n(e_i)$  which maximize it subject to constraint. Why does  $\exp(-E/T)$  already maximize  $n(e_i)$  so to speak? We noted above that since one sums over  $n$  in the photon case, there is no need to bring in any sum over degeneracies in momentum. Thus a

“microstate” calculation of unweighted (in terms of momentum degeneracies)  $\exp(-E/T)$  yields the same result as maximized  $\ln$  number of degeneracies using average  $n(e_i)$ . We consider this in the next section.

### Information Theory and a Microstate Containing All Degeneracy Information

Consider the case of  $i$  possible outcomes with probabilities  $p(e_i)$ . If one performs  $N$  trials, where  $N$  is very large, one has a specific microstate with the probability:

$$\text{Probability( microstate with } N \text{ very large)} = \text{Product over } i \text{ } p(e_i)^{Np(e_i)} \quad ((11))$$

Even though ((11)) is the result for a particular microstate, its value holds for any microstate in the large  $N$  case. Thus all microstates have the same probability and they are permutations of each other. The probability of a particular microstate ((11)) is equivalent to:  $1/\{\text{number of permutations}\}$  where:

$$\text{Number of permutations} = N! / \text{Product } n(e_i)! \quad ((12))$$

Thus the probability of a single microstate without considering any degeneracies (i.e. any permutations) contains the information of the number of permutations.

We argue that for the photon case,  $\hbar w$  represents a single energy scenario, but one may still sum over  $n$  to create a microstate result like ((11)) which has all of the information one needs. One does not have to explicitly add in degeneracy considerations i.e. one does not need to change:

$$\exp(-(\hbar w - u)n/T) \text{ to: } \{n+g(w)-1\}! / \{n! (g(w)-1)!\} \exp(-(\hbar w - u)n/T) \quad ((13))$$

when calculating the grand partition function.

### Conclusion

In conclusion, we argue that there seem to be two different commonly used approaches in statistical mechanics. The first is based on a partition function and assigns weights  $\exp(-(E-uN)/T)$  to microstates. In such a case,  $E = \sum_i e_i n(i)$  and  $N = \sum_i n(i)$ . If one does not sum over  $e_i$  states as in the case of only one type of photon present i.e. one with energy  $\hbar w$ , one still sums over  $n$  values. For a given  $n$ , there is momentum degeneracy because the  $n$  photon vectors may point in different directions. For a greater  $n$  value, there should be more arrangements of photon vectors, yet the weight  $\exp(-(\hbar w - u)n/T)$  does not bring in these degeneracies at all and yet still yields the Bose-Einstein distribution. An alternative approach is to assume an initial set of unknown  $n(e_i)$  without considering any microstates. In such a case, one must manually identify the permutations in each single particle level  $e_i$  with degeneracy  $g_i$  due to the momentum vector pointing in different directions. Taking  $\ln$  and maximizing to ensure that the least information is present subject to constraints  $\sum_i$

$e^{i n(e)}$  and  $\sum_i n(e_i)$  again yields the Bose-Einstein distribution. It thus seems that the microstate  $\exp(-(E-uN)/T)$  weights must already have the extra information linked with degeneracy removed. For example, for the MB case, maximizing  $-\ln \{ N! / \text{Product over } i n(e_i)! \}$  subject to the constraint  $\sum_i e_i n(e_i)$  yields  $\exp(-e_i/T)$  so starting with this value, one already has the form of minimized information and does not need to bring in  $n(e_i)!$  again to show that particles are identical.

## References

1. Reif, F. Fundamentals of Statistical and Thermal Physics (McGraw Hill, 1965)
2. Huang, K. Statistical Mechanics (Wiley, 1987)