Two ways to derive the BE and FD distribution functions

and some other guidance on statistical physics

Andrew M. Steane

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

February 4, 2026

We will obtain the Bose-Einstein and Fermi-Dirac distribution functions, which are

$$\langle n_i \rangle = \frac{1}{\exp(\beta(\epsilon_i - \mu) \pm 1} \tag{1}$$

where the plus sign should be taken for Fermions and the minus sign for Bosons. Here $\langle n_i \rangle$ is the mean occupation number of a single-particle quantum state $|i\rangle$ whose energy is ϵ_i, in a perfect gas at chemical potential μ and temperature T, with $\beta = 1/k_{\rm B}T$. The total number of particles in the gas is

$$N = \sum_i n_i \simeq \sum_i \langle n_i \rangle \tag{2}$$

where the approximation is very good in practice.

In the following we first give a derivation of the above using the concept of maximisation of entropy, under given constraints which are incorporated by Lagrange multipliers. We then give a derivation using the grand partition function. Finally, we will point that there is also a method quite like the second one, but which uses the canonical (as opposed to grand canonical) ensemble.

Why bother with two different methods? The Lagrange multiplier method is conceptually and algebraically simpler and gives the correct answer, but it uses some approximations which may not be valid in certain cases such as very low temperature. The grand partition function method, while a lot more involved, does not require approximations.

In all the methods, we model a gas as a collection of N particles. A quantum given state of the whole gas is specified by stating how many particles are in each of the possible single-particle quantum states $|i\rangle$. Such a specification can be given by furnishing the set of *occupation numbers* $\{n_i\}$. We then have

$$N = \sum_i n_i \tag{3}$$

and the total energy of the gas is

$$U = \sum_i n_i \epsilon_i \tag{4}$$

where ϵ_i is the energy of the single-particle state $|i\rangle$.

Note, the states available to the system can be specified in terms of a set of occupation numbers, whether we are dealing with distinguishable particles or indistinguishable fermions or indistinguishable bosons. The differences between the three cases concern the number of microstates associated with each set of occupation numbers.

1 Microcanonical (Method by entropy and Lagrange multipliers)

1.1 Fermions

Consider a gas of identical fermions. In practice there will usually be some large amount of degeneracy g_r of each energy level ϵ_r, and if there is not we can group states of similar energy together so that there are g_r quantum states in the r'th group.

Now suppose we want to assign n_r particles to the r'th group. Since for Fermions there can be at most 1 particle per state, the number of ways of doing this is equal to the number of ways to choose n_r items out of g_r items (without regard to ordering), which is

$$W_r = \frac{g_r!}{n_r!(g_r - n_r)!} \tag{5}$$

Note that for fermions the n_i can only be zero or one, but the n_r can be large when the g_r are large.

The number of ways of independently assigning particles to all the groups is

$$W = \Pi_r W_r. \tag{6}$$

Hence

$$\ln W = \sum_r \ln W_r \simeq \sum_r g_r \ln g_r - n_r \ln n_r - (g_r - n_r) \ln(g_r - n_r) \tag{7}$$

where we used Stirling's approximation, which here is valid when

$$g_r \gg 1, \qquad n_r \gg 1, \qquad (g_r - n_r) \gg 1. \tag{8}$$

(In practice we can usually ensure the first two conditions hold, but the third condition may not hold at very low temperatures; we shall return to this point.)

Now the entropy is proportional to $\ln W$ and we want to find a set of n_r which will maximise the entropy subject to the constraints (3),(4), which here take the form

$$N = \sum_r n_r = \text{const}, \qquad U = \sum_r n_r \epsilon_r = \text{const}. \tag{9}$$

From (7) we have

$$\frac{\partial \ln W}{\partial n_r} = \ln(g_r - n_r) - \ln n_r = \ln\left(\frac{g_r}{n_r} - 1\right). \tag{10}$$

The maximum of $\ln W$ is found when

$$\mathrm{d}\left(\ln \mathrm{W} + \alpha \mathrm{N} - \beta \mathrm{U}\right) = 0 \tag{11}$$

where α and β are Lagrange multipliers (the minus sign is taken so that later on it will be found that $\beta = 1/k_{\mathrm{B}}T$). Thus we have

$$\sum_r \left(\frac{g_r}{n_r} - 1\right) + \alpha - \beta\epsilon_r)\mathrm{dn_r} = 0 \tag{12}$$

and furthermore all the n_r can vary independently in this equation (the Lagrange multiplier method), therefore each term is itself zero:

$$\ln\left(\frac{g_r}{n_r} - 1\right) + \alpha - \beta\epsilon_r = 0. \tag{13}$$

Solving for n_r, we obtain

$$n_r = \frac{g_r}{\exp(-\alpha + \beta\epsilon_r) + 1}. \tag{14}$$

This is the number of fermions in a group of g_r quantum states, each having energy ϵ_r, in conditions of thermal equilibrium, in a gas of identical fermions. The number per quantum state is therefore

$$n_i = \frac{n_r}{g_r} = \frac{1}{e^{\beta(\epsilon_i - \mu)} + 1} \tag{15}$$

where we introduced $\mu = \alpha/\beta$. This μ is the chemical potential, a fact that can be deduced by using the results to deduce how much energy is required to add a particle to the system. Equation (15) is the Fermi-Dirac distribution function (compare with equation (1)).

1.2 Bosons

We shall treat now a perfect gas of bosons. As for the previous case, we suppose the quantum states are in groups of g_r at energy ϵ_r, and we will assign n_r particles to the r'th group. How many ways are there to do this for bosons? To answer this question, consider a line of n_r identical blobs and $g_r - 1$

Figure 1: Partitioning a set of objects. The illustration shows a case with $n_r = 10$, $g_r = 5$. We want to count ways of placing the n_r objects in g_r 'pigeonholes' when the objects are mutually indistinguishable (and there is no restriction on the number in any pigeonhole). The diagram shows one possible arrangement: it has $\{3, 2, 0, 1, 4\}$ particles allocated to each of the 5 states. By rearranging the $n_r + g_r - 1$ items (circles and lines) in the diagram we shall get other assignments whenever the rearrangement does not merely permute the circles among themselves or the lines among themselves.

partitions separating them into subsets (Fig. 1). In this way we arrive at g_r subsets, so we can see this as a way to assign particles to quantum states: each subset in our partitioning is associated with one quantum state. Clearly there are $(n_r + g_r - 1)!$ ways to arrange the blobs and partitions. But the associated quantum state of the whole set of particles will be the same if the blobs are rearranged among themselves, or if the partitions are rearranged among themselves. This reduces the total count by $n_r!(g_r - 1)!$ so we find

$$W_r = \frac{(n_r + g_r - 1)!}{n_r!(g_r - 1)!} \tag{16}$$

and therefore

$$\ln W_r \simeq (n_r + g_r)\ln(n_r + g_r) - n_r \ln n_r - g_r \ln g_r \tag{17}$$

where we used Stirling's approximation and took $g_r \gg 1$. Hence

$$\frac{\partial \ln W_r}{\partial n_r} = \ln(n_r + g_r) - \ln(n_r) = \ln\left(\frac{g_r}{n_r} + 1\right). \tag{18}$$

Notice how similar this is to equation (10). The rest of the argument goes exactly as it did for fermions, only with this sign change in the 1 term, and the result is

$$n_i = \frac{n_r}{g_r} = \frac{1}{e^{\beta(\epsilon_i - \mu)} - 1}. \tag{19}$$

This is the Bose-Einstein distribution function.

The reader should familiarize themselves with these two distribution functions; see Figures 2 and 3.

2 Grand Canonical ensemble

2.1 Probability of a microstate

In the grand canonical method we consider a system able to exchange both energy and particles with a large reservoir, and in equilibrium with that reservoir. In this case the system and the reservoir

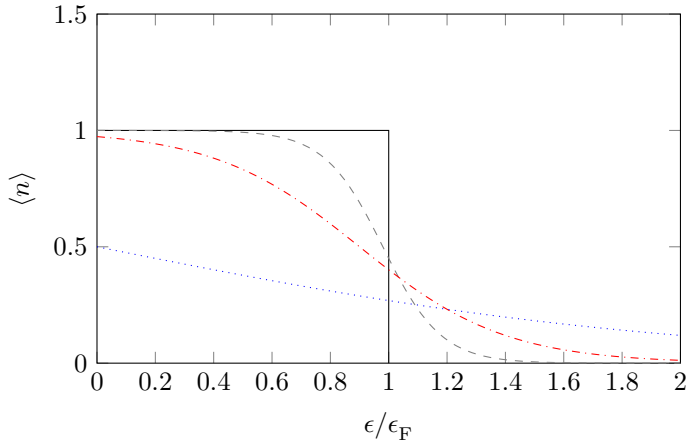


Figure 2: The Fermi-Dirac distribution function at four example temperatures. We plot the mean occupation of a state of energy ϵ, as a function of $\epsilon/\epsilon_{\mathrm{F}}$ where ϵ_{F} is the Fermi energy (the value of μ at $T = 0$). Full curve: $T = 0$; gray dashed: $k_{\mathrm{B}}T = 0.1\,\epsilon_{\mathrm{F}}$; red dash-dot: $k_{\mathrm{B}}T = 0.25\,\epsilon_{\mathrm{F}}$; blue dots: $k_{\mathrm{B}}T = \epsilon_{\mathrm{F}}$. In order to plot the function at any given T, one must allow for the fact that μ falls below ϵ_{F} as T increases from zero. One can locate μ on the graph of $\langle n(\epsilon) \rangle$ by observing that $\langle n \rangle = 0.5$ when $\epsilon = \mu$. (At higher temperatures still, $\langle n \rangle$ falls below 0.5 and μ is negative.)

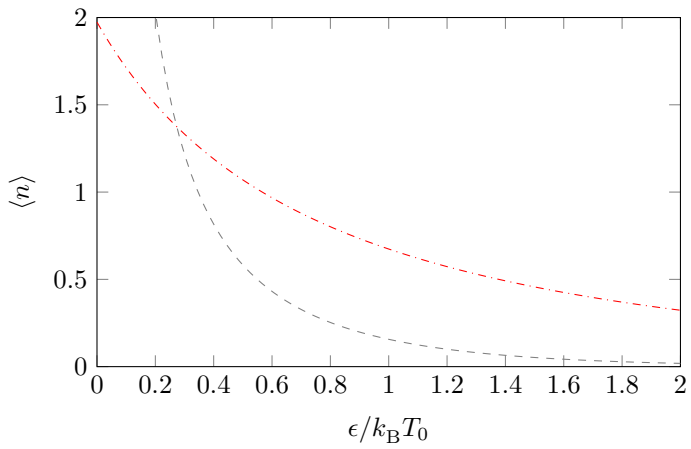
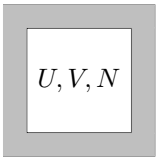
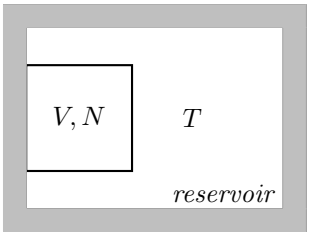


Figure 3: The Bose-Einstein distribution function at two example temperatures. We plot the mean occupation of a state of energy ϵ, as a function of $\epsilon/k_{\mathrm{B}}T_0$ where T_0 is the BEC transition temperature. The curve for $T = 0$ would be a delta-function spike at $\epsilon = 0$. Gray dashed: $T = 0.5\,T_0$; red dash-dot: $k_{\mathrm{B}}T = 2\,T_0$. The former case (T below T_0) has μ very close to zero, with the result that the ground state occupation is of order N, a large number. The latter case (T above T_0) has $\mu \simeq -0.82 k_{\mathrm{B}}T_0$ with the result that none of the quantum states has a macroscopic population.



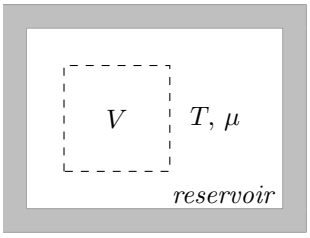
"microcanonical"
U, V, N fixed
$W = W(U, V, N)$

$S = k_{\mathrm{B}} \ln W$



"canonical"
T, V, N fixed
U fluctuates
$Z = Z(T, V, N)$.

$F = -k_{\mathrm{B}} T \ln Z$



"grand canonical"
T, V, μ fixed
U and N fluctuate
$\mathcal{Z} = \mathcal{Z}(T, V, \mu)$

$\Omega = -k_{\mathrm{B}} T \ln \mathcal{Z}$

Figure 4: The three main approaches to statistical thermal physics. The shading indicates isolation. The Helmholtz function $F = U - TS$. The grand potential $\Omega = F - \mu N$. Each approach leads to an expression for a thermodynamic potential in terms of its natural variables, and therefore to all other thermodynamic information.

share the same temperature and the same chemical potential. In the basic reasoning to follow, $T_{\rm R}$ and $\mu_{\rm R}$ refer to properties of the reservoir in the first instance, but we can employ the thermodynamic result $T = T_{\rm R}$ and $\mu = \mu_{\rm R}$. We also introduce $\beta = 1/k_{\rm B}T$ as usual.

The starting point of the method is to find a formula for the probability that the system will be found in a quantum state (i.e. a single microstate) having N particles and energy E. The formula is

$$p(N, E) \propto e^{\beta(\mu N - E)}. \tag{20}$$

We will now derive this. (Note, if there is more than one microstate having the given N and E this is ok; we will be finding the probability for the system to be found in any one microstate.)

The system and reservoir together form a composite system which is isolated, and therefore has fixed volume V_0, energy U_0 and particle number N_0. In the following, we use a subscript zero to refer to properties of the composite, a subscript R to refer to properties of the reservoir, and no subscript for the system. Let $W_{\rm R}(U_{\rm R}, N_{\rm R})$ be the number of microstates of the reservoir for given macroscopic reservoir properties U_R, N_R. We shall consider a *single microstate* of the system, one in which the system has energy N particles and energy E. Whenever the system has N, E the reservoir must have $N_0 - N$ and particles and energy $U_0 - E$, so the number of microstates of the reservoir is

$$W_{\rm R}(U_0 - E,\, N_0 - N). \tag{21}$$

Since we are considering a single microstate of the system, this is also the number of microstates available to the entire isolated composite system. On the assumption of equal a priori probabilities in an isolated system, it follows that the probability for this situation to arise is

$$\begin{aligned} p(N, E) &= \text{const.}\, W_{\rm R}(U_0 - E,\, N_0 - N) &(22)\\ &= \text{const.}\, e^{S_{\rm R}(U_0 - E,\, N_0 - N)}. &(23)\end{aligned}$$

Now expand the entropy in a Taylor series:

$$S_{\rm R} = S_{\rm R}(U_0, N_0) - \frac{\partial S_{\rm R}}{\partial U_{\rm R}} E - \frac{\partial S_{\rm R}}{\partial N_{\rm R}} N + \frac{1}{2} \frac{\partial^2 S_{\rm R}}{\partial U_{\rm R}^2} E^2 + \ldots \tag{24}$$

where we have written just one of the second-order terms but we will consider all of them in a moment. Next, from the fundamental thermodynamic relation $\mathrm{dU} = \mathrm{TdS} - \mathrm{pdV} + \mu\mathrm{dN}$ one deduces

$$\frac{\partial S_{\rm R}}{\partial U_{\rm R}} = \frac{1}{T_{\rm R}}, \qquad \frac{\partial S_{\rm R}}{\partial N_{\rm R}} = -\frac{\mu_{\rm R}}{T_{\rm R}}. \tag{25}$$

Substituting these in (24) we have

$$S_{\rm R} = S_{\rm R}(U_0, N_0) - \frac{E}{T_{\rm R}} + \frac{\mu_{\rm R} N}{T_{\rm R}} - \frac{1}{2T_{\rm R}^2} \frac{\partial T_{\rm R}}{\partial U_{\rm R}} E^2 + \ldots \tag{26}$$

Now we need to argue why it is legitimate to drop the second- and higher-order terms in this expression. We observe that $(\partial U_{\rm R}/\partial T_{\rm R})$ is the constant-volume heat capacity of the reservoir, $C_{\rm R}$. The ratio of this second-order to the first-order term in the Taylor expansion is therefore

$$\frac{E}{2T_{\rm R} C_{\rm R}} \ll 1. \tag{27}$$

The important point is that *we can always make the reservoir large enough to ensure that this quantity is small, and furthermore it tends to zero in the thermodynamic limit for the reservoir.* (To see that the ratio is indeed small, observe that for typical states of a system in thermal equilibrium, E is at most a few times $Nk_{\mathrm{B}}T$ whereas C_{R} can be of order $N_{\mathrm{R}}k_{\mathrm{B}}T$, and $N_{\mathrm{R}} \gg N$.) Similar arguments apply to the other second-order terms which we have not explicitly written, and to the third- and higher-order terms. Therefore in the thermodynamic limit of the reservoir, we obtain

$$S_{\mathrm{R}} = S_{\mathrm{R}}(U_0, N_0) - \frac{E}{T_{\mathrm{R}}} + \frac{\mu_{\mathrm{R}} N}{T_{\mathrm{R}}}. \tag{28}$$

By substituting this into (23), and using $\mu = \mu_{\mathrm{R}}$ and $T = T_{\mathrm{R}}$, equation (20) follows. Note that although the method proceeded via a Taylor series expansion, we have argued that the result is exact in the thermodynamic limit of the reservoir.

A remark on the reasoning via $S_{\mathrm{R}}(U_0 - E,\, N_0 - N)$.
In the above reasoning it is crucial to consider the second- and higher-order terms, not just the first-order terms. Here is why. If we had set out from (22) we could have expanded W_{R} in a Taylor series, and had we kept only the first order terms in that expansion, we would have found a *different* formula for $p(N, E)$ (in fact one then gets a linear approximation to $p(N, E)$ so it is not completely wrong, but it is not sufficiently accurate to be of any use). The point is that the correct result requires the *whole* Taylor series to be considered. We found that the linear approximation to S_{R} is correct in the thermodynamic limit: the entropy really does vary linearly with changes in U_{R} and N_{R} of the order of E and N, whereas the number of microstates W_{R} does not (except for unusually small values of E and N).

A reminder:

Question: why is it $\beta(\mu N_j - E_j)$ in the grand partition function?

Answer: because the fundamental relation of thermodynamics has $\mathrm{dS} = \frac{1}{\mathrm{T}}\left(\mathrm{dE} - \mu \mathrm{dN} + \mathrm{pdV}\right)$ and we are treating a rigid system.

2.2 Partition function

Having obtained the crucial equation (20) we can now proceed to

$$p(N_j, E_j) = \frac{1}{\mathcal{Z}} e^{\beta(\mu N_j - E_j)} \tag{29}$$

where

$$\mathcal{Z} = \sum_j e^{\beta(\mu N_j - E_j)}. \tag{30}$$

In these equations, j is a label for a quantum state $|j\rangle$ in which the system has N_j particles and energy E_j, and the sum is over all available states. The symbol $\mathcal{Z}$ is called the *grand canonical partition*

function (or just *grand partition function*). Note, there can be many quantum states having the same N_j or E_j or both; (e.g. it may be that $N_1 = N_2 = N_3$ and $E_6 = E_7$ or whatever); each quantum state is included as one term in the sum.

We can use the grand partition function to find the average number of particles, for the average is given by

$$\langle N \rangle = \sum_j N_j p_j \tag{31}$$

(where $p_j \equiv p(N_j, E_j)$) and we have

$$\frac{\partial \ln \mathcal{Z}}{\partial \mu} = \frac{1}{\mathcal{Z}} \sum_j \beta N_j e^{\beta(\mu N_j - E_j)} = \beta \sum_j N_j p_j = \beta \langle N \rangle \tag{32}$$

hence

$$\langle N \rangle = \frac{1}{\beta} \frac{\partial \ln \mathcal{Z}}{\partial \mu}. \tag{33}$$

This result will be useful in finding the Fermi-Dirac and Bose-Einstein distributions. We just need to find an expression for $\mathcal{Z}$ in terms of macroscopic properties of the system (that is, T, V, μ).

So far the reasoning could apply to any type of system. From now on we focus on a perfect gas. In this context a gas is called 'perfect' or 'ideal' when it can be treated as a collection of particles which do not interact, except just enough to allow them to come to thermal equilibrium. This situation is called *weakly interacting*. It means we assume the quantum states and their energies are not influenced by the movement of particles between states. In this situation each microstate $|j\rangle$ of the entire system can be associated with a unique set of *occupation numbers* $\{n_i\}$. We next need to decide whether this mapping is one-to-one or many-to-one:

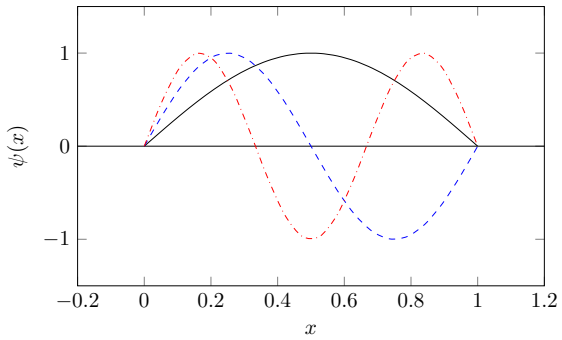
- Distinguishable particles: many different $|j\rangle$ can have the same $\{n_i\}$.
- Indistinguishable particles: there is only one quantum state $|j\rangle$ for each set of occupation numbers $\{n_i\}$.

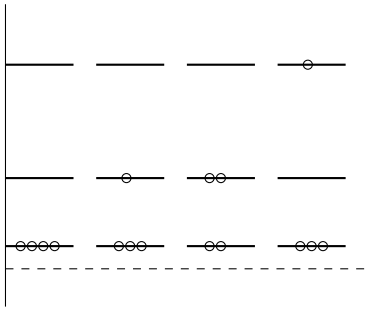
From now on we shall limit our discussion to the indistinguishable case: there is a one-to-one correspondence between each quantum state $|j\rangle$ of the entire system and a set of *occupation numbers* $\{n_i\}$. The occupation numbers say how many particles are in the first single-particle state, how many in the second single-particle state, and so on. This is an important basic concept so we repeat it:

$$\begin{array}{ccc} |j\rangle & = & |n_1\rangle\,|n_2\rangle\,|n_3\rangle \cdots \\ j & \overset{\text{1-to-1}}{\longleftrightarrow} & \{n_1, n_2, n_3, \ldots\} \end{array} \tag{34}$$

and

$$N_j = \sum_i n_i, \qquad E_j = \sum_i n_i \epsilon_i \tag{35}$$





j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\{n_1, n_2, n_3\}$	400	310	220	301	130	211	040	121	202	031	112	022	103	013	004
E_j	4	7	10	12	13	15	16	18	20	21	23	26	28	31	36

Figure 5: A worked example of the relationship between states and occupation numbers. The plot shows the first three wavefunctions for states of a single particle in a one-dimensional box (infinite square well). Underneath are listed the 15 possible states of the whole system when there are 4 indistinguishable bosons, if we imagine they are restricted to just these three wavefunctions. The diagram on the right is a pictorial representation of the first four cases. Because the particles are indistinguishable, there is a single quantum state associated with each set of occupation numbers: see eqn (34).

where ϵ_i are the energies of the single-particle states. See Figure 5 for an example.

By using these expressions in (30) we find

$$\mathcal{Z} = \sum_{n_1}\sum_{n_2}\sum_{n_3}\cdots\sum_{n_M} \exp\left(\beta\sum_{i=1}^{M}(\mu-\epsilon_i)n_i\right) \tag{36}$$

where M is the number of single-particle states (which may be infinite). It is important to note that by summing over each n_i we account correctly for each and every state $|j\rangle$, without double-counting or missing any out. (Had we been dealing with distinguishable particles we would be massively under-counting and the above expression would be quite wrong.) Next we use the fact that an exponential of a sum is equal to a product of exponentials:

$$e^{\beta\mu(n_1+n_2+\cdots)-\beta(n_1\epsilon_1+n_2\epsilon_2+\cdots)} = e^{\beta(\mu-\epsilon_1)n_1}e^{\beta(\mu-\epsilon_2)n_2}\cdots . \tag{37}$$

Employing this in (36) we have:

$$\mathcal{Z} = \sum_{n_1}\sum_{n_2}\cdots\sum_{n_M} e^{\beta(\mu-\epsilon_1)n_1}e^{\beta(\mu-\epsilon_2)n_2}\cdots \tag{38}$$

$$= \left[\sum_{n_1} e^{\beta(\mu-\epsilon_1)n_1}\right]\left[\sum_{n_2} e^{\beta(\mu-\epsilon_2)n_2}\right]\cdots . \tag{39}$$

States, modes, occupation numbers
The word 'state' is much employed in statistical physics, and in the case of multiple particles it can be confusing. For example, one may say there are n_1 particles in the first state, n_2 in the second, and so on, and at the end of such a list one has described just one state of the entire system! For this reason I think it better to say there are 'n_1 particles in the first *single-particle state*, n_2 in the second,' etc. An even better terminology is 'n_1 particles in the first *mode*, n_2 in the second,' etc. Here the term 'mode' expresses the fact that strictly speaking we are concerned with quantum field theory and the wave vectors $\mathbf{k}$ used to describe the available single-particle 'states' are really describing the plane-wave modes of a bosonic or fermionic field. The 'occupation number' is the degree of excitation of the mode.

This is a remarkable result and the reader is encouraged to convince themselves that the multiple sum does indeed factorize in this way (for example, consider the simple case $(A^0B^0 + A^0B^1 + A^1B^0 + A^1B^1) = (A^0 + A^1)(B^0 + B^1)$).

We now have

$$\mathcal{Z} = \prod_{i=1}^{M} \mathcal{Z}_i, \tag{40}$$

$$\text{where} \qquad \mathcal{Z}_i = \sum_{n_i} e^{\beta(\mu-\epsilon_i)n_i}. \tag{41}$$

It is tempting to 'read' the quantity $\mathcal{Z}_i$ as another partition function, and so it is, as we now show. What we need to do is realise that the overall probability of the entire system state j can itself be written as a product:

$$p_j \equiv p(N_j, E_j) = p(n_1, n_2, \cdots) = \prod_{i=1}^{M} p_i(n_i) \tag{42}$$

where

$$p_i(n_i) = \frac{1}{\mathcal{Z}_i} e^{\beta(\mu-\epsilon_i)n_i}. \tag{43}$$

This factorization of the probability follows from the fact that the particles are acting independently and furthermore the occupation numbers $n_1, n_2, \cdots$ are not constrained to add up to any particular number. In these conditions the statistical distribution for any given single-particle state is independent of the others. (The reader is also invited to notice that (43) is indeed the unique formula which, upon being substituted into (42), will yield (29) for all values of μ, β.)

We have now arrived at a convenient way to proceed. We can focus on one particular single-particle state: the one labelled by i. This state is itself a (somewhat abstract) kind of 'system' which can have energy and particles. When it has n_i particles it has energy $n_i\epsilon_i$. But how can a single-particle state have lots of particles? The answer is one of terminology. The phrase 'single particle state' refers to a

wavefunction which can be that of a single particle. For a gas of many particles, the situation is that zero or more of the particles may have that wavefunction at any given time. But there is another, and I think clearer, way of describing the situation. I find the particle terminology here confusing. Instead one can adopt field terminology. We say the physical system is the i'th field-mode, and n_i is the degree of excitation of that mode. This way of putting it into words makes it clear what physical system we are referring to (a field mode), and also the reasoning is closer to the quantum field theory which underlies the whole discussion, so that one is reassured that no mistake has been made.

There are just two steps remaining to obtain the distribution functions. First, by an argument essentially the same as that which led to (33) we find

$$\langle n_i \rangle = \frac{1}{\beta} \frac{\partial \ln \mathcal{Z}_i}{\partial \mu}. \tag{44}$$

Hence if we can find an expression for $\mathcal{Z}_i$ then $\langle n_i \rangle$ will follow. But an expression for $\mathcal{Z}_i$ is easy to find, for (41) is a geometric series! (It is $\sum_{n_i} z_i^{n_i}$ where $z_i = \exp(\beta(\mu - \epsilon_i))$. In the case of bosons the series runs from $n_i = 0$ to $n_i = \infty$ and we find

$$\mathcal{Z}_i = \frac{1}{1 - e^{\beta(\mu - \epsilon_i)}} \qquad \text{bosons.} \tag{45}$$

In the case of fermions there are just two terms $n_i = 0$ and $n_i = 1$ and we find

$$\mathcal{Z}_i = 1 + e^{\beta(\mu - \epsilon_i)} \qquad \text{fermions.} \tag{46}$$

By substituting these expressions into (44) we obtain

$$\langle n_i \rangle = \begin{cases} \dfrac{1}{e^{\beta(\epsilon_i - \mu)} + 1} & \text{fermions} \\ \\ \dfrac{1}{e^{\beta(\epsilon_i - \mu)} - 1} & \text{bosons} \end{cases} \tag{47}$$

2.3 Summary: the essence of the grand ensemble method

Let us now summarise.

The method of grand canonical ensemble to find distribution functions can be divided into four main steps.

Step 1. The first step is to show that the probability distribution over energy and particle number, in thermal equilibrium, is

$$p(N, E) \propto e^{\beta(\mu N - E)}. \tag{48}$$

Step 2. The second step is to introduce

$$\mathcal{Z} = \sum_j e^{\beta(\mu N_j - E_j)}. \tag{49}$$

and it is worth noticing that the mean value of N can be found from

$$\langle N\rangle = \frac{1}{\beta}\frac{\partial \ln \mathcal{Z}}{\partial \mu}. \tag{50}$$

Step 3. Next one shows that *for a set of indistinguishable particles which behave independently*, (or, clearer in my opinion: *for a set of distinguishable field modes*):

$$\mathcal{Z} = \prod_{i=1}^{M} \mathcal{Z}_i, \tag{51}$$

$$\text{where} \qquad \mathcal{Z}_i = \sum_{n_i} e^{\beta(\mu-\epsilon_i)n_i}. \tag{52}$$

in which n_i is the number of particles in the i'th single-particle state (if you want to talk about particles), or, equally, n_i is the degree of excitation of the i'th field mode (if you prefer to talk about fields) and M is the number of field modes. One also argues that the probability to find n_i particles in the i'th single-particle state is

$$p_i(n_i) = \frac{1}{\mathcal{Z}_i} e^{\beta(\mu-\epsilon_i)n_i} \tag{53}$$

and therefore

$$\langle n_i\rangle = \frac{1}{\beta}\frac{\partial \ln \mathcal{Z}_i}{\partial \mu}. \tag{54}$$

Step 4. Finally, one notices that the sum in (52) is simply a geometric series so the sum can be performed. In the case of fermions the range of n_i is 0 to 1; in the case of bosons the range is 0 to ∞. The expressions for $\mathcal{Z}_i$ are thus obtained, and by differentiating (as indicated in (54)) the Fermi-Dirac and Bose-Einstein distributions follow.

If one is in a hurry one might begin at step 3, asserting (53) without proof. Or, for more thoroughness, one could produce an argument for (53) without bothering with the N-particle case, by arguing that the i'th field mode is itself a physical system to which quantum and statistical arguments can be applied.

3 Canonical ensemble

This section can be omitted at first reading.

For systems of distinguishable particles, the canonical ensemble is simpler to use than the grand ensemble, and this is why it is normally introduced before the latter in discussions of statistical mechanics. However in the case of a system of indistinguishable particles the canonical approach is difficult when we cannot ignore the indistinguishability. This is the case when occupation numbers are not small compared to 1, and this situation arises at low temperature.

Having said that, the canonical method is not entirely forfeit. The argument given in section 2 sometimes causes students to think that in order to treat a perfect gas of identical particles we need the grand canonical method. This is not so. If the gas has a fixed number of particles then one can equally well use the canonical or the microcanonical approaches. Furthermore, the case of thermal radiation can also be treated by a canonical approach since one does not need to model the radiation as a gas of particles (with unconstrained total number of particles); one can model it instead (and arguably more clearly) as a set of modes, each of which behaves like a quantum harmonic oscillator.

In this section we will consider a gas with a fixed number of particles, N. The canonical partition function is

$$Z_N = \sum_j e^{-\beta E_j} \tag{55}$$

where j labels the quantum states of the entire gas. The relationship between states and occupation numbers (34) still holds, but now the sets of occupation numbers all add to the same total N, so (35) is replaced by

$$\sum_i n_i = N, \qquad E_j = \sum_i n_i \epsilon_i. \tag{56}$$

We therefore have

$$Z_N = \sum_{\{n_i\}} e^{-\beta \sum_i n_i \epsilon_i} \tag{57}$$

where the notation in the first sum is a shorthand for summing over all sets of occupation numbers that satisfy (56).

Now the mean occupation number of state $|k\rangle$ is

$$\langle n_k \rangle = \sum_j n_k p_j = \frac{1}{Z_N} \sum_j n_k e^{-\beta E_j} \tag{58}$$

where the sum over j is the sum over all sets $\{n_i\}$ satisfying (56). The reader should note that this mean value is indeed given by summing over j, not i. We use the probability p_j of each state $|j\rangle$ of the whole gas and multiply by the value of n_k for that state, then take the sum. Now observe from (57) that

$$\frac{\partial Z_N}{\partial n_k} = \sum_j -\beta n_k e^{-\beta E_j}. \tag{59}$$

Hence we find

$$\langle n_k \rangle = -\frac{1}{\beta} \frac{\partial \ln Z_N}{\partial \epsilon_k}. \tag{60}$$

This will allow us to find the distributions functions, once we have an expression for Z_N.

Unfortunately, finding Z_N is difficult because the sum over $\{n_i\}$ in (57) is hard to do, and cannot be done exactly (in any simple analytic way). But a neat mathematical trick (or observation or relevant fact, if you prefer) makes it all easy again.

We now introduce a mathematical trick. Rather than trying to calculate Z_N directly, we will make it our aim to calculate the following quantity:

$$X \equiv \sum_{N=0}^{\infty} e^{-\alpha N} Z_N \tag{61}$$

where α is a constant introduced for the purpose of writing this X. The idea behind the trick is that we can calculate X (as we show in a moment) and furthermore we can estimate Z_N to sufficient accuracy by taking it that the sum determining X is dominated by one term (the largest one, obviously). That is, we shall be taking

$$X \simeq e^{-\alpha N^*} Z_{N^*} \tag{62}$$

where N^* is the value of N for which

$$\frac{\partial}{\partial N}\left(e^{-\alpha N} Z_N\right) = 0. \tag{63}$$

Using just one term (equation (62)) to estimate an infinite sum (61) can be surprising at first. In fact this method does not estimate Z_N itself very well, but it gives a very good estimate of $\ln Z_N$ and all the thermodynamic properties depend on $\ln Z_N$, not Z_N directly. Section A shows how the step from (61) to (62) can be justified.

Now we need to carefully consider the sums involved in X. We have

$$X = \sum_{N=0}^{\infty} \sum_{\{n_i\}} (\text{blah}) \tag{64}$$

where (blah) refers to whatever we are summing (it is $\exp(-\alpha N - \beta E_j)$ in this example). The $N = 0$ term in this double sum includes just those sets $\{n_i\}$ with all $n_i = 0$. The $N = 1$ term includes all those cases where one of the n_i is 1 and the others are zero. The $N = 2$ term includes all those where $\sum_i n_i = 2$, and so on. By such reasoning one finds that the combination of sums in the definition of X is exactly matched by another combination of sums:

$$\sum_{N=0}^{\infty} \sum_{\{n_i\}} (\text{blah}) = \sum_{n_1} \sum_{n_1} \sum_{n_2} \cdots \sum_{n_M} (\text{blah}) \tag{65}$$

It follows that, if we take $\alpha = -\beta\mu$, then our X is none other than the grand partition function $\mathcal{Z}$ given by (36)!

We can now use the previous discussion of the grand partition function, and thus obtain

$$X = \mathcal{Z} = \prod_i \left(1 \pm e^{\beta(\mu - \epsilon_i)}\right)^{\pm 1} \tag{66}$$

(equations (40),(45),(46)) where the plus sign is taken for fermions the minus sign for bosons. Hence

$$Z_N = e^{\alpha N} \prod_i \left(1 \pm e^{\beta(\mu-\epsilon_i)}\right)^{\pm 1} \tag{67}$$

where we used (62), taking $N = N^*$ (one chooses the value of α so as to bring about this agreement between N and N^*).

Finally, by substituting our expression for Z_N into the formula for $\langle n_k \rangle$, eqn. (60), we have

$$\langle n_k \rangle = -\frac{1}{\beta}\frac{\partial}{\partial \epsilon_k}\left[\alpha N \pm \sum_i \ln\left(1 \pm z_i\right)\right] \tag{68}$$

where

$$z_i = e^{-(\alpha+\beta\epsilon_i)}. \tag{69}$$

(This z_i is introduced purely in order to make the next bit of algebra easier to read.) Hence

$$\langle n_k \rangle \quad = \quad -\frac{1}{\beta}\left[\frac{\partial \alpha}{\partial \epsilon_k} N + \sum_i \frac{1}{1 \pm z_i}\frac{\partial z_i}{\partial \epsilon_k}\right]. \tag{70}$$

Now

$$\frac{\partial z_i}{\partial \epsilon_k} = -\left(\frac{\partial \alpha}{\partial \epsilon_k} + \beta \frac{\partial \epsilon_i}{\partial \epsilon_k}\right) z_i, \tag{71}$$

therefore

$$\langle n_k \rangle \quad = \quad -\frac{1}{\beta}\frac{\partial \alpha}{\partial \epsilon_k}\left\{N - \sum_i \frac{z_i}{1 \pm z_i}\right\} + \sum_i \frac{z_i}{1 \pm z_i}\frac{\partial \epsilon_i}{\partial \epsilon_k}. \tag{72}$$

We will show in a moment that the term in the curly bracket is zero. Also, we have $(\partial \epsilon_i / \partial \epsilon_k) = 1$ when $i = k$ and it equals zero when $i \neq k$, hence it acts to pick out one term from the sum, giving

$$\langle n_k \rangle = \frac{z_k}{1 \pm z_k} = \frac{1}{(z_k)^{-1} \pm 1} = \frac{1}{e^{\alpha+\beta\epsilon_k} \pm 1} \tag{73}$$

where the plus sign is to be taken for fermions, the minus sign for bosons. After substituting $\alpha = -\beta\mu$ we have the same distribution functions as before (c.f. eqn (1)).

To complete the argument, we should tidy up two points. First, the argument that the curly bracket in (72) evaluates to zero is simply that the sum in that bracket can now be recognized as $\sum_i \langle n_i \rangle$ which must evaluate to $\langle N \rangle$ and furthermore this must equal N in the canonical ensemble. The other point is the relationship between α and μ. We argued that α is chosen so as to make the maximum term in (61) be the one where N in the sum is equal to the number of particles actually in our system. We then noted that as long as $\alpha = -\beta\mu$ then $X = \mathcal{Z}$. To check the overall consistency we now need

to confirm that this relationship between α and μ is the one which the canonical partition function predicts. Using the expression for Z_N, eqn (67), we can find μ by standard methods. We have the fundamental thermodynamic relation

$$\begin{aligned} \mathrm{dU} &= T\mathrm{dS} - \mathrm{pdV} + \mu\mathrm{dN}, \\ \Rightarrow \qquad \mathrm{dF} &= -S\mathrm{dT} - \mathrm{pdV} + \mu\mathrm{dN}, \end{aligned} \tag{74}$$

and therefore

$$\mu = \left.\frac{\partial F}{\partial N}\right|_{T,V}. \tag{75}$$

Using $F = -k_\mathrm{B}T \ln Z_N$ one finds by an easy differentiation that $\mu = -k_\mathrm{B}T\alpha$, as required.

A Method to approximate an integral over a peak

In statistical thermal physics one often encounters sums of the form

$$Z = \sum_i g(x_i) e^{-\alpha x_i} \tag{76}$$

for some function $g(x)$ and constant α. Usually the function $g(x)$ is monotonically increasing, but not exponentially so. In this case the terms in the sum will grow with x for small x, and decrease with x for large x. In other words the function $(g(x)e^{-\alpha x})$ has a peak at some x (and no other peaks).

A good estimate of the sum can often be obtained by approximating it by an integral. With this in mind, let's consider the general task of evaluating any integral of the form

$$z = \int_0^\infty f(x)\,\mathrm{dx} \tag{77}$$

where the function $f(x)$ has a single peak at some positive x. The idea is to fit a gaussian curve to the peak, and then use the integral under the gaussian curve as our approximation to z (see Fig. 6).

If the fitted gaussian curve has a height $f(x_0)$ and standard deviation Δx then the result is

$$z \simeq f(x_0)\Delta x\sqrt{2\pi}. \tag{78}$$

The value of x_0 is obtained by finding the location of the peak: it is given by the value of x at which $\mathrm{df/dx} = 0$. A suitable value for Δx is obtained as follows.

Let $y = \ln f$. Then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f'/f, \qquad \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -(f'/f)^2 + f''/f. \tag{79}$$

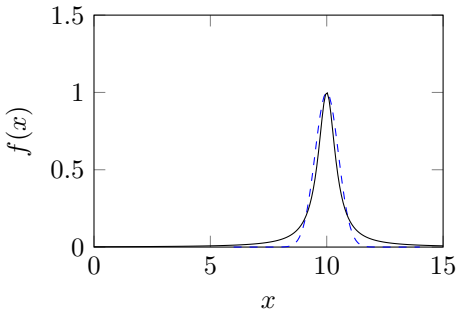


Figure 6: The full line shows some peaked function $f(x)$. We would like to estimate the area under the curve. It can be done by fitting a gaussian curve (shown dashed) to the peak, and calculating the area under that.

It follows that the Taylor expansion of y around the peak at $x = x_0$, up to second order, is

$$y \simeq y(x_0) + \frac{1}{2}(x - x_0)^2 \frac{f''}{f}. \tag{80}$$

Hence we find, for values of x near the peak,

$$f(x) \simeq f(x_0) \exp(-(x - x_0)/2\Delta x^2) \tag{81}$$

where

$$\Delta x = \sqrt{-1/y''} = \sqrt{-f/f''} \tag{82}$$

(and this is to be evaluated at $x = x_0$). This furnishes the information we need to evaluate our approximation to the integral, (78).

A.1 Stirling approximation

As an example of the method, let's consider the factorial function:

$$n! = \int_0^\infty x^n e^{-x} \mathrm{dx}. \tag{83}$$

This equation is exact. But let's get an approximate evaluation of the integral using the above method. We have $f(x) = x^n e^{-x}$ and therefore $y(x) = n \ln x - x$. Then $\mathrm{dy/dx} = 0$ at $x = n$ and at this value of x, $y'' = -1/n$. Hence we find $\Delta x = \sqrt{n}$. The method yields

$$n! \simeq f(n)\Delta x\sqrt{2\pi} = n^n e^{-n}\sqrt{2\pi n}. \tag{84}$$

This is called *Stirling's approximation*. It is a fairly good approximation to $n!$, and it yields a very good approximation to $\ln(n!)$:

$$\ln(n!) \simeq n \ln n - n + (1/2) \ln(2\pi n). \tag{85}$$

A.2 The precision of approximation of logarithms

For $n \gg 1$ the final term in Stirling's approximation for $\ln(n!)$ (85) is entirely negligible compared to the others, so then we have

$$\ln(n!) \simeq n \ln n - n \tag{86}$$

which is how the approximation is often used in practice. And this illustrates a rather remarkable fact. The latter result asserts that $n! \simeq n^n e^{-n}$, in other words

$$\int x^n e^{-x} \mathrm{dx} \simeq (\mathrm{x^n e^{-x}} \text{ evaluated at } \mathrm{x = n}). \tag{87}$$

In this expression, *the integral over all of x is being estimated by the maximum value of the integrand.* But one should be cautious of such an estimate. In fact by dropping the $\ln(2\pi n)$ term in (85) we have that our expression for $\ln(n!)$ is now inaccurate by about $\ln(\sqrt{2\pi n})$ and therefore the expression on the right of (87) is inaccurate by a factor $\sqrt{2\pi n}$, which is a large factor for $n \gg 1$.

What is going on? The above is an example of a general fact about very large numbers and logarithms. To make the point, a numerical example suffices. Suppose we have some quantity W whose true value is $W = \exp(10^{20})$ (note, this is an extremely large number!). Suppose now that we have a method to estimate W, but unfortunately our estimate is wrong by a factor one million. Ordinarily one would think that an estimate wrong by a factor one million is entirely useless for scientific purposes. But it is not so! For, what if the only way W influences the physical effects we wish to calculate is through its logarithm. Then we have that the true value is

$$\ln W = 10^{20} \tag{88}$$

and our faulty estimate gives

$$\ln(10^6 W) = 6 \ln 10 + 10^{20}. \tag{89}$$

In this example, therefore, the imprecision in our estimate of $\ln W$ is just $6 \ln 10 \simeq 13.8$ for a quantity whose magnitude is 10^{20}. This is an accuracy of about one part in 10^{19}, so very precise indeed.

We are now ready to justify the step from (61) to (62), in which a sum was approximated by a single term in the sum. One first argues that the sum can be estimated by an integral, and then one applies the reasoning leading to (78). Finally, one drops the $\Delta x \sqrt{2\pi}$ factor, which is valid in the case where one is really interested in the logarithm of the sum in the end, as opposed to the sum itself. In the case of (61), Δx corresponds to ΔN, the size of the thermal fluctuations in the particle number, and the result is accurate when this is small compared to N. In a wide range of circumstances one finds $\Delta N \simeq \sqrt{N}$ which is indeed small compared to N when $N \gg 1$.