

Outline Notes: Lectures 1-7

1 Introduction (Lecture 1)

1.1 Macroscopic and microscopic descriptions

Microstate

- Defined by the microscopic variables: e.g., p_i, x_i for each particle.
- Each microstate corresponds to an (orthogonal) quantum state of the whole system, e.g., energy eigenstates.
- A microstate provides a full, detailed description of the state of the system.

Macrostate

- Defined by macroscopic (e.g., thermodynamic) variables: e.g., (U, V, N) .
- A macrostate is an incomplete description of the state of the system.
- Each macrostate may have a (very) large number of corresponding microstates.

The statistical weight Ω

- The *statistical weight* of a macrostate is the number of microstates in the macrostate
- We denote it by Ω .
- The statistical weight is sometimes called the *multiplicity* — they are the same thing. You may even sometimes see it called the *thermodynamic probability*. However, it's not really a probability — it's usually a very large integer!

1.2 Single-particle states

Consider a system of N identical particles. The complete wavefunction for a system of N particles is a function of every coordinate of every particle: $\psi(r_1, r_2, \dots, r_N, \dots)$. For N particles, the number of interaction terms in the Schrödinger equation $\sim N^2$. A full computation of these terms becomes completely infeasible even for relatively small numbers of particles. We therefore make a key simplification and describe the system in terms of single-particle states.

- We assume particles are only *weakly interacting* so that particles can be treated independently of each other.
- We then consider a single particle and calculate the possible states for that particle. For example, for a container of gas, we might obtain them by solving the Schrödinger equation for one particle of the gas. These states are the *single-particle states*, ψ_j for $j = 0, 1, 2, \dots$

- For a system of identical independent particles, each particle can be described using the same set of single-particle states.
- We can then define a microstate of the whole system by specifying the single-particle state for every particle. We can write a microstate as $\Psi = (j_1, j_2, \dots, j_N)$ which would specify a microstate where particle 1 is in ψ_{j_1} , particle 2 in ψ_{j_2} , etc.

NOTE: Do not confuse “single-particle state” with “microstate”: the set of single-particle states describes the possible states for just one particle; a microstate specifies the state of every particle in the system. (For the special case that the system has only one particle, then the microstates are simply the single-particle states.)

1.3 Distinguishable and indistinguishable particles

In statistical mechanics it matters whether particles are distinguishable or indistinguishable. The number of distinct microstates we can count depends on whether particles are distinguishable or not.

Distinguishable: atoms in a solid

Indistinguishable: gases, liquids, electrons in metals, photons

In quantum mechanics, all atoms or molecules of the same species, electrons, photons, etc. are fundamentally identical. So how do we get distinguishable particles?

It's not really the particles that are distinguishable or indistinguishable: nearly always they are indistinguishable. One copper atom is identical to another. Its constituents (electrons, nucleons, etc.) are fundamentally identical and indistinguishable. Moreover, since their wavefunctions cannot be completely distinct, there is no way even in principle of tagging or following individual identical particles through the system.

The real difference between distinguishable and indistinguishable is in the nature of the single-particle states that we use to describe the system.

In a gas we use one (large) set of single-particle states, occupying the whole volume of the gas. This set of states is essentially shared between all the particles of the gas. The fundamental identicalness of the particles means there is no way of distinguishing between these two descriptions:

- *particle A in state 1 and particle B in state 2;*
- *particle A in state 2 and particle B in state 1.*

If particle A and particle B are fundamentally the same, the two descriptions are actually of the same microstate. The confusion arises because we have tried to label the particles. It is much clearer if we say:

- *state 1 is occupied by a particle and state 2 is occupied by a particle.*

In a solid, we describe the system using a separate set of single-particle states for each lattice site. The sets of single-particle states at each lattice site may be identical to each other (one copper atom in a lump of copper is the same as another) however each lattice site essentially has its own copy. We can therefore distinguish between these two descriptions:

- the atom at site A in state 1 and the atom at site B in state 2;
- the atom at site A in state 2 and the atom at site B in state 1.

In this case we say the particles are distinguishable, but what we really mean is the lattice sites of the particles are distinguishable.

1.4 Ensembles

Our goal in statistical physics is to obtain macroscopic properties of a system from the microscopic description. There are typically an extremely large number of microstates to deal with. Moreover, the microstate of the system will be constantly changing with the microscopic motion of the system. We cannot hope to know which microstate the system is in at any given time. We need a statistical description of the system in terms of the available microstates.

Definition An *ensemble* is a probability distribution for the microstates of a system.

- The ensemble provides a statistical description of the system based on the microscopic states.
- In equilibrium, the ensemble does not change with time. This does not mean that the microscopic state of the system is not changing — that may be changing continuously, for example as molecules in a gas move and collide. However the probability of a finding the system in any particular microstate does not vary with time.

If we know the appropriate ensemble for the system, we can use it to calculate macroscopic properties and behaviour of the system. The question then is what is the appropriate ensemble to use.

NOTE: An alternative definition of an ensemble is a very large assembly of identical systems that collectively represent the microscopic states of the system. We can obtain the probabilities by dividing the number of systems in a particular microstate by the total number of systems in the assembly. The two descriptions therefore give essentially the same information, but it is sometimes a useful construct to think of a large assembly of identical systems.

2 Isolated systems

2.1 The microcanonical ensemble

We'll begin with an isolated system of distinguishable particles in equilibrium. Isolated systems are relatively simple to describe and specify in terms of macroscopic quantities. For an isolated system U , V and N are fixed. These are all well-defined for the macrostate. For a reasonably sized system the number of microstates consistent with such a description is very large. We need to decide what probabilities to ascribe to the different microstates, in other words what should we use for the ensemble. We have no further information about the microstates so we use the **fundamental postulate** of statistical mechanics:

- *In an isolated system, every accessible microstate is equally likely.*

Justification: The fundamental postulate is really just a statement of ignorance. We have only specified U , V , and N , so any microstate that corresponds to these macroscopic variables is allowed. We have no other information to expect one particular microstate more than any other. That is *exactly* what we mean by “equally likely”.

Note, not all microstates are *accessible*. Microstates have to be consistent with the external constraints on U , N , V , i.e., they have to be in the macrostate.

We can then write a probability distribution: the probability of a microstate is

$$p = 1/\Omega$$

for microstates that are in the macrostate and zero for all other microstates. This defines the **microcanonical ensemble**. The microcanonical ensemble is applicable for an isolated system at equilibrium.

Note that the larger Ω is, the less ordered the system is from a statistical view.

2.2 The Boltzmann entropy

We define the Boltzmann entropy of the system as

$$S_B = k_B \ln \Omega$$

where k_B is a constant — Boltzmann's constant.

At this stage, we treat this as a postulate.

2.2.1 Einstein's argument

Consider a system AB comprising two independent systems A and B with numbers of microstates given by Ω_A and Ω_B . The total number of microstates of the combined system AB is

$$\Omega_{AB} = \Omega_A \times \Omega_B.$$

For each microstate in A we can have any microstate in B — each distinct combination is a separate microstate of the system AB. However, entropy is an extensive quantity, so the entropy of the combined system is the sum of the entropies of the two systems

$$S_{AB} = S_A + S_B.$$

Therefore if S is a function of Ω , then it must satisfy $S(\Omega_A \times \Omega_B) = S(\Omega_A) + S(\Omega_B)$ which requires that it be of the form

$$S = k \ln \Omega.$$

2.3 Configurations and distributions

We are rarely interested in the microscopic details of exactly which particle is in which single-particle state. Indeed for all but the smallest of systems, that would be an information overload. Of far more use is the distribution of particles among the single particle states, or perhaps a distribution of particles among energy levels. We will refer to this as a *configuration* or sometimes as a *distribution*.

Consider a system of N identical distinguishable particles each of which is described by the same set of single-particle states ψ_j for $j = 0, 1, 2, \dots$

- We write a microstate of the system by specifying the single-particle state for each particle, i.e., (j_1, j_2, \dots, j_N) .
- The number of particles that are in the j^{th} single-particle state we refer to as the occupancy of that state, which we denote n_j .
- The list of all the occupancies of the single-particle states, $[n_0, n_1, \dots, n_j, \dots]$ specifies a *configuration*. I.e., it is a distribution of the particles among the single-particle states. We denote this configuration \mathbf{n} .
- Not all configurations have the same number of microstates: some may have many more than others.
- Example: microstates $(1, 3, 4, 3, 0, 1, 1, 1)$ and $(0, 3, 1, 3, 1, 4, 1, 1)$ are just two of the many microstates that belong to the configuration $[1, 4, 0, 2, 1, 0, \dots]$, whereas the configuration $[0, 0, 8, 0, \dots]$ has only one microstate $(2, 2, 2, 2, 2, 2, 2, 2)$.
- The statistical weight of a configuration \mathbf{n} is the number of microstates that belong to the configuration, which we will denote by $\Omega_{\mathbf{n}}$.
- The probability of a particular configuration \mathbf{n} is

$$p_{\mathbf{n}} = \Omega_{\mathbf{n}} / \Omega \quad \text{with} \quad \Omega = \sum_{\mathbf{n}} \Omega_{\mathbf{n}}$$

where Ω is the statistical weight of the macrostate, which is equal to the sum of the statistical weights of all the configurations of the macrostate.

A macrostate will typically have a very large number of possible configurations and each configuration may in turn have a very large number of microstates. However, since different configurations may contain very different numbers of microstates, some configurations will be much more likely than others.

2.4 The equilibrium configuration

The most likely configuration is the one that has the largest number of microstates. This is a direct consequence of the fundamental postulate. We refer to this as the *equilibrium configuration*. (It is sometimes called the *equilibrium state* or the *equilibrium distribution*.)

One of our aims will be to determine the equilibrium configuration for a system.

Note that the system is not always in the equilibrium configuration. Firstly, not all systems are in equilibrium. However, even at “thermodynamic equilibrium” the system is not always in this exact equilibrium state configuration. The system will evolve, e.g., collisions in gas, exchange of vibration energy in a solid, etc. According to the fundamental postulate all microstates are equally likely — this includes those states that may correspond to situations far from equilibrium. In other words there will be fluctuations.

Question: Is the most-likely state a good representation of the system in equilibrium?

- Short answer: yes.
- Slightly longer answer: only when we consider configurations that are “almost” the same as the most likely. The sum of the Ω_n for these states dominates the total available microstates. The probability of significant departures from the equilibrium state becomes negligible for sufficiently large systems. (see Problem Sheet 1)

3 Maximizing entropy (Lecture 2)

3.1 Example: entropy for distinguishable spin-half particles

Consider an isolated system consisting of an assembly of N identical distinguishable spin-half particles. There are just two single particle states ψ_j for $j = 1, 2$ corresponding to spin-up and spin-down states respectively. Both states have the same energy. A system microstate is defined by specifying the spin state for each particle, (j_1, j_2, \dots, j_N) . Each particle can be in either the spin-up or spin-down state, and so the total number of possible microstates is 2^N . All these microstates are equally likely. The Boltzmann entropy of the system is

$$S_B = k_B \log \Omega = k_B \ln 2^N = k_B N \ln 2.$$

We may ask how many of the particles are in each of the spin states. The configurations are specified by the occupancy numbers of the single-particle states $[n_1, n_2]$, where n_1 is the number with spin-up and n_2 is the number with spin-down. For N particles, the possible configurations are $\mathbf{n} = [m, N-m]$ for $m = 0..N$. The likelihood of each configuration depends on its statistical weight. The statistical weight for the configuration $[m, N-m]$ is

$$\Omega_{\mathbf{n}} = \frac{N!}{m!(N-m)!}.$$

The corresponding entropy for the configuration is

$$S_{B\mathbf{n}} = k_B \ln \Omega_{\mathbf{n}} = k_B [\ln N! - \ln m! - \ln(N-m)!].$$

Using Stirling's approximation for large numbers, this gives

$$S_{B\mathbf{n}} = k_B [N \ln N - m \ln m - (N-m) \ln(N-m)].$$

The equilibrium configuration is that with the largest entropy (which is equivalent to largest statistical weight). This occurs for $m = N/2$, giving the entropy for the equilibrium configuration as

$$S_{B\mathbf{n}} = k_B (N \ln N - N \ln N/2) = k_B N \ln 2.$$

Hence for large numbers the entropy for the equilibrium configuration is very close to the total entropy (and to Stirling's approximation they happen to be the same for this example.)

3.2 The Boltzmann distribution

We will now use the maximum entropy approach to derive the Boltzmann distribution for an assembly of identical distinguishable particles. We will start from the microcanonical ensemble.

Consider an isolated system that contains an assembly of identical distinguishable particles, for example a solid comprising N identical atoms at fixed lattice sites. Since the system is isolated, its total energy U and its volume V are fixed. Each atom in the assembly is described by the set of single-particle states ψ_j with energy levels ϵ_j . We wish to find the equilibrium configuration, i.e., the most likely distribution of the atoms among the single-particle states. From the fundamental postulate, the microcanonical ensemble is appropriate i.e., all accessible microstates are equally probable. However, not every possible microstate is accessible, only those that have total energy U . We therefore need to find the configuration with the largest statistical weight subject to constraints on energy.

3.2.1 Maximizing the entropy

Each configuration is specified by a list of occupancy numbers of the single-particle states, $[n_0, n_1, \dots]$, where n_0 is the number of particles in the ground state, n_1 is the number in the first state, and so on. There may be many distinct ways of distributing the particles among the single-particle states that result in the same configuration. Counting all the different arrangements, i.e., the microstates, for a configuration gives the statistical weight of that configuration, Ω_n .

Question: How many microstates are there for a configuration $[n_0, n_1, \dots]$?

Answer:

$$\Omega_n = \frac{N!}{\prod_j n_j!}$$

To find the most-likely configuration, we maximize $\ln \Omega_n$ subject to constraints,

$$\sum_j n_j = N, \quad \sum_j \epsilon_j n_j = U$$

Using Lagrange's method of undetermined multipliers,

$$\frac{\partial}{\partial n_j} \left[\ln \Omega_n - \alpha \sum_j n_j - \beta \sum_j \epsilon_j n_j \right] = 0 \quad \text{for all } j.$$

This requires that

$$\frac{\partial}{\partial n_j} \ln \Omega_n - \alpha - \beta \epsilon_j = 0 \quad \text{for all } j.$$

Substituting for Ω_n ,

$$-\frac{\partial}{\partial n_j} \ln n_j! - \alpha - \beta \epsilon_j = 0 \quad \text{for all } j.$$

For large n_j (using Stirling's approximation),

$$\ln n_j + \alpha + \beta \epsilon_j = 0 \quad \text{for every } j,$$

giving

$$n_j = e^{-\alpha - \beta \epsilon_j}.$$

This is the most likely distribution of particles among the single-particle states. It is the Boltzmann distribution.

The Lagrange multipliers α and β , which are yet to be determined, are associated with the constraints on the total number of particles N and the total energy U .

3.2.2 Temperature — significance of β (Lecture 3)

The fundamental equation of thermodynamics is

$$TdS = dU + dW.$$

For a system with constant V , $dW = 0$, so we have $dU = TdS$. This leads to the thermodynamic definition of temperature,

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_{V,N}.$$

In the statistical description we began with the Boltzmann definition of entropy,

$$S_B = k_B \ln \Omega.$$

We found the equilibrium configuration by maximizing the entropy of the configuration, $k_B \ln \Omega_n$. Now consider changing the internal energy of the system, keeping V and N fixed. This will give a new equilibrium configuration. The change in internal energy is

$$\delta U = \sum_j \epsilon_j \delta n_j.$$

Keeping V fixed means that the energy levels ϵ_j do not change. The change in the Boltzmann entropy is

$$\delta S_B = k_B \sum_j \frac{\partial \ln \Omega_n}{\partial n_j} \delta n_j.$$

But for the equilibrium configuration,

$$\frac{\partial \ln \Omega_n}{\partial n_j} = \alpha + \beta \epsilon_j.$$

Therefore,

$$\begin{aligned} \delta S_B &= k_B \alpha \sum_j \delta n_j + k_B \beta \sum_j \epsilon_j \delta n_j \\ &= k_B \beta \delta U \end{aligned}$$

(using $\sum \delta n_j = 0$ since N is fixed) and so

$$\beta = \frac{1}{k_B} \left(\frac{\partial S_B}{\partial U} \right)_{V,N}.$$

This agrees exactly with the thermodynamic definition of temperature and we see that β is related to the thermodynamic temperature,

$$\beta = 1/k_B T.$$

Thus we have arrived at a microscopic/statistical description of temperature. The distribution of particles among states varies exponentially with the energy of the state; the temperature defines the energy scale of this distribution. In terms of the temperature, the distribution is

$$n_j = e^{-\alpha} e^{-\epsilon_j/k_B T} \propto e^{-\epsilon_j/k_B T}.$$

3.2.3 The number of particles — significance of α

The Lagrange multiplier constant α was introduced to constrain total number of particles.

$$N = \sum_j n_j \quad \Rightarrow \quad N = e^{-\alpha} \sum_j e^{-\beta \epsilon_j} \quad \Rightarrow \quad e^{-\alpha} = \frac{N}{\sum_j e^{-\beta \epsilon_j}}.$$

Thus the value of α normalizes the distribution to the total number of particles.

The summation in the above will turn out to be very important in deriving thermodynamic properties of the system. It is called the *partition function* which we will denote by Z ,

$$Z = \sum_j e^{-\beta \epsilon_j}$$

(Z from the German *Zustandssumme* — sum over states). We shall return to the partition function later. In terms of the partition function,

$$e^{-\alpha} = N/Z.$$

Finally we have for the Boltzmann distribution,

$$n_j = \frac{N}{Z} e^{-\beta \epsilon_j} = \frac{N}{Z} e^{-\epsilon_j / k_B T}$$

where $\beta = 1/k_B T$ and Z is the partition function. (Note: some treatments define $\beta = -1/k_B T$. The choice is arbitrary depending on the sign of the multipliers in the Lagrangian. It does not affect the physics.)

The probability that a particle is in the j^{th} state is then given by

$$p_j = \frac{e^{-\beta \epsilon_j}}{Z}.$$

3.3 Zeroth law

Consider two isolated systems, which need not be the same, having internal energies U_1 and U_2 . For each system, at equilibrium their Boltzmann entropies are maximal, and we have

$$\beta_1 = \frac{1}{k_B} \left(\frac{\partial S_{B1}}{\partial U_1} \right)_{V,N}, \quad \beta_2 = \frac{1}{k_B} \left(\frac{\partial S_{B2}}{\partial U_2} \right)_{V,N}.$$

The entropy of the combined system is

$$S_{B12} = S_{B1} + S_{B2}.$$

Now consider placing the systems in weak thermal contact so that energy can transfer between them. The change in entropy due to energy δU transferring from system 1 to system 2 is

$$\delta S_{B12} = \left[\left(\frac{\partial S_{B2}}{\partial U_2} \right)_{V,N} - \left(\frac{\partial S_{B1}}{\partial U_1} \right)_{V,N} \right] \delta U = (\beta_2 k_B - \beta_1 k_B) \delta U.$$

However, at equilibrium, the entropy of the combined system is maximal, so $\delta S_{B12} = 0$ which implies $\beta_1 = \beta_2$. Therefore systems in thermal equilibrium with each other must have the same value of β .

4 Systems in heat baths (Lecture 4)

4.1 The canonical ensemble

We have already derived the Boltzmann distribution using an isolated system comprising an assembly of distinguishable particles. The nature of the particles, their internal details, was not specified. The only requirement was that they were in weak thermal contact with each other. Each particle was effectively in a heat bath provided by the rest of the assembly. Since the particle nature was not specified, the derivation could be applied equally to any distinguishable particles: atoms, molecules, or indeed arbitrary systems (replacing “particle” with “system”.) In fact the Boltzmann distribution is quite general and describes any system in contact with a heat bath.

To see this more directly, consider an isolated system AB consisting of two components, a system A and a heat bath B, that are in weak thermal contact with each other. A heat bath can be any body with a heat capacity very large compared to the system that it serves so that its temperature is unaffected by thermal exchange with the system. The heat bath's internal details do not matter. For a given internal energy U , the number of microstates of the heat bath will be very large. From definition of Boltzmann entropy we have for the statistical weight

$$\Omega(U) = e^{S/k_B}.$$

Consider a small change in the internal energy of the heat bath δU . From the definition of temperature, the entropy change is

$$\delta S = \left(\frac{\partial S}{\partial U} \right)_V \delta U = \frac{1}{T} \delta U,$$

which gives a change in the statistical weight,

$$\Omega(U + \delta U) = \Omega(U) e^{\delta S/k_B} = \Omega(U) e^{\delta U/k_B T}.$$

Thus, the statistical weight of a bath increases exponentially with its internal energy.

Returning to the combined system AB, its microstates are formed by combining each microstate of A with each microstate of the heat bath B that results in a total energy U . For a particular microstate of A with energy ϵ_j , the number of available microstates of the heat bath is given by $\Omega(U - \epsilon_j)$. Summing these weights for all microstates of A gives the total statistical weight of the combined system,

$$\Omega_{AB} = \sum_j \Omega(U - \epsilon_j) = \Omega(U) \sum_j e^{-\epsilon_j/k_B T}.$$

Since all these microstates of AB are equally likely, the probability for a particular microstate of A that has energy ϵ_j is simply the statistical weight for that microstate divided by the total statistical weight,

$$p_j = \frac{\Omega(U - \epsilon_j)}{\Omega_{AB}} = \frac{e^{-\epsilon_j/k_B T}}{\sum_j e^{-\epsilon_j/k_B T}} = \frac{1}{Z} e^{-\epsilon_j/k_B T},$$

where Z is the partition function for system A. This is the probability distribution for the microstates of a system in contact with a heat bath. It is of course the Boltzmann distribution again. This probability distribution defines the canonical ensemble.

Summary In the canonical ensemble the probability of each microstate of the system is given by

$$p_j = \frac{1}{Z} e^{-\epsilon_j/k_B T}, \quad \text{where} \quad Z = \sum_j e^{-\epsilon_j/k_B T}.$$

Z is the partition function for the system. The canonical ensemble applies to a system in thermal equilibrium with a heat bath of temperature T .

4.2 General expression for entropy (Gibbs entropy)

Consider a large assembly of N identical, independent systems distributed among the system microstates according to a probability distribution p_j . (The p_j defines the ensemble.) The systems are distinguishable. The total number of systems in the j^{th} state is given by Np_j . Then the number of microstates for this distribution is

$$\Omega = \frac{N!}{\prod_j (Np_j)!}.$$

Using Stirling's formula,

$$\ln \Omega = N \ln N - N - \sum_j [Np_j \ln(Np_j) - Np_j].$$

Noting that $\sum p_j = 1$, this becomes

$$\ln \Omega = -N \sum_j p_j \ln p_j.$$

The total Boltzmann entropy of the assembly is then

$$S_B = k_B \ln \Omega = -Nk_B \sum_j p_j \ln p_j.$$

Assuming the entropy to be extensive, the entropy of a single system is

$$S = -k_B \sum_j p_j \ln p_j.$$

This is the Gibbs entropy and is a generalization of the Boltzmann entropy. For the case of an isolated system described by the microcanonical ensemble, the probabilities of all the microstates are equal and this reduces to the Boltzmann definition of entropy.

We can use this to calculate the entropy of a system in contact with a heat bath (i.e., a system described by the canonical ensemble).

4.2.1 Example: 2-state system

Consider a system with two states with energies ϵ_1 and ϵ_2 in a heat bath at temperature T . The partition function is given by

$$Z = e^{-\beta\epsilon_1} + e^{-\beta\epsilon_2}$$

and the probabilities for the microstates are

$$p_1 = \frac{1}{1 + e^{-\beta\Delta\epsilon}}, \quad p_2 = \frac{e^{-\beta\Delta\epsilon}}{1 + e^{-\beta\Delta\epsilon}},$$

where $\Delta\epsilon = \epsilon_2 - \epsilon_1$.

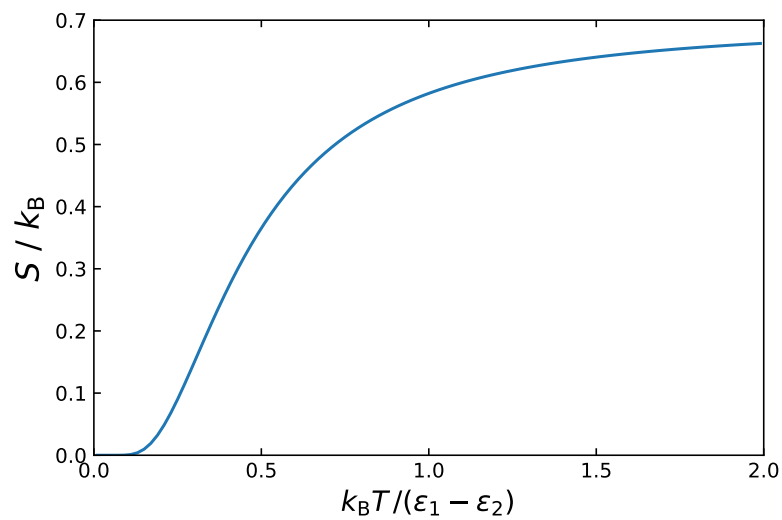


Figure 1: Gibbs entropy for a two-state system

Behaviour:

$$\text{For } T \rightarrow 0, \quad p_1 = 1, \quad p_2 = 0 \quad \Rightarrow \quad S = -k_B(1 \log 1 + 0 \log 0) = 0.$$

$$\text{For } T \rightarrow \infty, \quad p_1 = \frac{1}{2}, \quad p_2 = \frac{1}{2} \quad \Rightarrow \quad S = -k_B\left(\frac{1}{2} \log \frac{1}{2} + \frac{1}{2} \log \frac{1}{2}\right) = k_B \log 2.$$

4.3 Link to thermodynamics (Lecture 5)

We shall see that the partition function provides a link to the thermodynamical description of the system. We begin by showing that the mean energy of the system (i.e., the expectation value) can be expressed straightforwardly in terms of the partition function. The mean energy is

$$\langle E \rangle = \sum_j p_j \epsilon_j = \frac{1}{Z} \sum_j \epsilon_j e^{-\beta \epsilon_j}.$$

But note that

$$\frac{\partial Z}{\partial \beta} = \frac{\partial}{\partial \beta} \sum_j e^{-\beta \epsilon_j} = - \sum_j \epsilon_j e^{-\beta \epsilon_j}.$$

Therefore,

$$\langle E \rangle = - \frac{1}{Z} \frac{\partial Z}{\partial \beta} = - \frac{\partial \ln Z}{\partial \beta}.$$

We identify the mean energy of the system with its internal energy U . Therefore we have

$$U = \langle E \rangle = - \frac{1}{Z} \frac{\partial Z}{\partial \beta} = - \frac{\partial \ln Z}{\partial \beta}.$$

In terms of T ($T = 1/k_B \beta$), we have

$$U = - \frac{\partial \ln Z}{\partial T} \frac{dT}{d\beta} = k_B T^2 \frac{\partial \ln Z}{\partial T}.$$

4.4 Fluctuations

The canonical ensemble gives a probability distribution over states of the system with different energies. Implicit in this description is that the energy will fluctuate. Since the system can exchange heat with its surroundings, the energy is no longer fixed. Following the same approach used to calculate the internal energy, we can calculate the energy fluctuations. The variance of the energy is

$$(\Delta E)^2 = \langle (E - \langle E \rangle)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2.$$

But,

$$\langle E^2 \rangle = \sum_j \epsilon_j^2 p_j = \frac{1}{Z} \sum_j \epsilon_j^2 e^{-\beta \epsilon_j} = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2}$$

and

$$\langle E \rangle = - \frac{1}{Z} \frac{\partial Z}{\partial \beta}.$$

However, differentiating $\ln Z$ twice we have,

$$\frac{\partial^2 \ln Z}{\partial \beta^2} = \frac{\partial}{\partial \beta} \frac{1}{Z} \frac{\partial Z}{\partial \beta} = \underbrace{\frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2}}_{\langle E^2 \rangle} - \underbrace{\frac{1}{Z^2} \left(\frac{\partial Z}{\partial \beta} \right)^2}_{\langle E \rangle^2}.$$

Therefore the variance of the energy is

$$(\Delta E)^2 = \frac{\partial^2 \ln Z}{\partial \beta^2} = -\frac{\partial U}{\partial \beta}.$$

In terms of temperature T ,

$$(\Delta E)^2 = -\frac{\partial U}{\partial T} \frac{dT}{d\beta} = k_B T^2 \frac{\partial U}{\partial T} = k_B T^2 C_V$$

since $\partial U / \partial T$ is the heat capacity of the system at constant volume, C_V . Thus the standard deviation of the energy fluctuations is

$$\Delta E = \sqrt{\frac{C_V}{k_B}} k_B T.$$

Note that the heat capacity is an extensive quantity. Classically, from equipartition we expect $k_B/2$ per degree of freedom in the heat capacity. For a system with n degrees of freedom, this gives $C_V = nk_B/2$, and we have

$$\Delta E = \sqrt{\frac{n}{2}} k_B T \quad \Rightarrow \quad \Delta E = \frac{C_V T}{\sqrt{n/2}} \sim \frac{\langle E \rangle}{\sqrt{n}}.$$

For example, for an ideal monatomic gas, with three degrees of freedom per molecule, $\Delta E = \sqrt{(3N/2)} k_B T$, where N is the number of molecules. For a macroscopic system (i.e., very large n), the relative size of the fluctuations is very small.

Question: How do we reconcile the small size of these fluctuations with the Boltzmann distribution which tells us the probability of any microstate is proportional to $e^{-\beta E}$?

The answer lies in the way the number of available microstates varies with energy. Recall that for a system of N oscillators, the statistical weight varied as $\Omega \sim E^N$. More generally, for a large system with a large number of degrees of freedom, the number of microstates is an extremely rapidly increasing function of energy. The probability of finding the system at a particular energy is proportional to the number of microstates available at that energy times the Boltzmann factor, so

$$p(E) \sim E^N e^{-E/k_B T}.$$

This product of the rapidly increasing E^N factor with the rapidly decreasing Boltzmann factor has a sharp maximum at $E \approx N k_B T$ — the larger N is, the sharper the peak. Conversely, for a small system with a small number of degrees of freedom, such as a single oscillator, the fluctuations can be large.

4.5 Free energy — the Helmholtz function

To relate the statistical description to the macroscopic thermodynamic description we need to be able to calculate macroscopic thermodynamic quantities. In thermodynamics, the properties of the system can be calculated from the thermodynamic potential functions. The free energy of the system allows us to calculate any other thermodynamic quantities. For a fixed volume system in contact with a heat bath, the free energy is the Helmholtz function

$$F = U - TS.$$

At equilibrium the free energy is a minimum.

In terms of the statistical description, using the Gibbs entropy we have for TS ,

$$TS = -k_B T \sum_j p_j \ln p_j.$$

But,

$$p_j = \frac{e^{-\epsilon_j/k_B T}}{Z} \Rightarrow \ln p_j = -\frac{\epsilon_j}{k_B T} - \ln Z.$$

Therefore,

$$\begin{aligned} TS &= \sum_j p_j \epsilon_j + k_B T \sum_j p_j \ln Z \\ &= U + k_B T \ln Z. \end{aligned}$$

Rearranging we get

$$F = U - TS = -k_B T \ln Z.$$

Thus we have found a simple expression for the Helmholtz free energy in terms of the statistical description. For the canonical ensemble this provides an explicit link between the statistical description and the thermodynamic description. We see that the partition function, which was originally introduced as a normalization factor plays a key role. Once we know the partition function for our system, we can in principle calculate any thermodynamic quantity via our expression for the Helmholtz function. This is sometimes referred to as the “Bridge equation” or the “Royal route”. For example,

$$S = - \left(\frac{\partial F}{\partial T} \right)_V, \quad P = - \left(\frac{\partial F}{\partial V} \right)_T.$$

If we know F in terms of the partition function, we can calculate internal energy,

$$U = F + TS = F - T \left(\frac{\partial F}{\partial T} \right)_V.$$

Using the statistical expression for the Helmholtz function, $F = -k_B T \ln Z$, gives

$$\begin{aligned} U &= -k_B T \ln Z + k_B T^2 \frac{\partial \ln Z}{\partial T} + k_B T \ln Z \\ \Rightarrow U &= k_B T^2 \frac{\partial \ln Z}{\partial T}, \end{aligned}$$

which is the same as the expression derived earlier from the probabilities p_j .

4.6 Partition function factorization (Lecture 6)

The partition function for a system described by the canonical ensemble is given by

$$Z = \sum_j e^{-\beta \epsilon_j},$$

where the summation is over all the microstates of the system (note: *microstates* not *energy levels*, which is important if there are degenerate states.) However, the number of microstates is typically very large and evaluating this summation for a complex system can be hard. In practice for systems of weakly interacting identical particles it is easier to calculate the partition function of a single particle first. Consider a single particle system with microstates Ψ_j and corresponding energy levels ϵ_j . The partition function for this single particle system is

$$Z_{(1)} = \sum_j e^{-\beta \epsilon_j}.$$

Now consider a system with two identical weakly interacting distinguishable particles. The microstates states are now $\Psi = \Psi_{j_1} \Psi_{j_2}$ for every combination of the two single-particle states and the energies are $\epsilon_{j_1} + \epsilon_{j_2}$. The partition function for this two-particle system is

$$Z_{(2)} = \sum_{j_1} \sum_{j_2} e^{-\beta(\epsilon_{j_1} + \epsilon_{j_2})} = \left(\sum_{j_1} e^{-\beta \epsilon_{j_1}} \right) \left(\sum_{j_2} e^{-\beta \epsilon_{j_2}} \right) = Z_{(1)}^2.$$

The partition function factorizes for the case of weakly interacting particles. Similarly for N particles, the partition function is

$$Z = Z_{(1)}^N.$$

Then, for the system of N particles

$$\ln Z = \ln Z_{(1)}^N = N \ln Z_{(1)}.$$

This means that $\ln Z$ is an *extensive* quantity for such a system. This is important as $\ln Z$ is a key link with thermodynamic quantities, e.g.,

$$F = -k_B T \ln Z = -N k_B T \ln Z_{(1)}, \quad U = k_B T^2 \frac{\partial \ln Z}{\partial T} = N k_B T^2 \frac{\partial \ln Z_{(1)}}{\partial T}.$$

Factorization of the partition function can also apply for composite systems comprised of weakly interacting components (so that the wavefunction factorizes). For example, consider a single diatomic molecule in a gas. The molecule has translational, vibrational and rotational components to its motion which can be treated as independent of each other giving states with the total energy

$$\epsilon = \epsilon_{\text{trans}} + \epsilon_{\text{vib}} + \epsilon_{\text{rot}}.$$

The partition function then factorizes

$$\begin{aligned} Z &= \sum_{j_{\text{trans}}} \sum_{j_{\text{vib}}} \sum_{j_{\text{rot}}} e^{-\beta(\epsilon_{j_{\text{trans}}} + \epsilon_{j_{\text{vib}}} + \epsilon_{j_{\text{rot}}})} = Z_{\text{trans}} Z_{\text{vib}} Z_{\text{rot}} \\ \Rightarrow \quad \ln Z &= \ln Z_{\text{trans}} + \ln Z_{\text{vib}} + \ln Z_{\text{rot}}, \end{aligned}$$

allowing us to treat the different components of the system separately.

4.7 Partition function: harmonic oscillators

Consider a 1-D quantum harmonic oscillator in thermal equilibrium at temperature T . For example, we can think of perhaps a single atom in a solid. The appropriate ensemble in this case is the canonical ensemble. The energy levels for a 1-D harmonic oscillator are

$$\epsilon_j = (j + \frac{1}{2})\hbar\omega \quad j = 0, 1, 2, \dots$$

The partition function is then

$$\begin{aligned} Z &= e^{-\frac{1}{2}\hbar\omega\beta} \sum_{j=0}^{\infty} e^{-j\hbar\omega\beta} \\ &= \frac{e^{-\frac{1}{2}\hbar\omega\beta}}{1 - e^{-\hbar\omega\beta}}, \end{aligned}$$

where we have used the standard summation $\sum_{j=0}^{\infty} x^j = 1/(1-x)$.

In terms of the partition function, the internal energy is given by

$$U = -\frac{\partial \ln Z}{\partial \beta}.$$

Taking the log of Z ,

$$\ln Z = -\frac{1}{2}\hbar\omega\beta - \ln(1 - e^{-\hbar\omega\beta})$$

and differentiating with respect to β , the internal energy is,

$$\begin{aligned} U &= \frac{1}{2}\hbar\omega + \frac{1}{1 - e^{-\hbar\omega\beta}}(-e^{-\hbar\omega\beta})(-\hbar\omega) \\ &= \left(\frac{1}{2} + \frac{1}{e^{\hbar\omega\beta} - 1} \right) \hbar\omega. \end{aligned}$$

Note that it is often easier to work in terms of β rather than T . This is the internal energy for one oscillator. For a solid comprising N identical distinguishable oscillators, the internal energy is N times this.

Behaviour at low temperature

$$\begin{aligned} \hbar\omega\beta \gg 1 &\Rightarrow \frac{1}{e^{\hbar\omega\beta} - 1} \ll 1 \\ &\Rightarrow U \approx \frac{1}{2}\hbar\omega. \end{aligned}$$

Behaviour at high temperature

$$\begin{aligned} \hbar\omega\beta \ll 1 &\Rightarrow e^{\hbar\omega\beta} \approx 1 + \hbar\omega\beta + \dots \\ \Rightarrow U &\approx \left(\frac{1}{2} + \frac{1}{\hbar\omega\beta} \right) \hbar\omega = \frac{\hbar\omega}{2} + k_B T \approx k_B T. \end{aligned}$$

Note: this agrees with the classical result for high temperature.

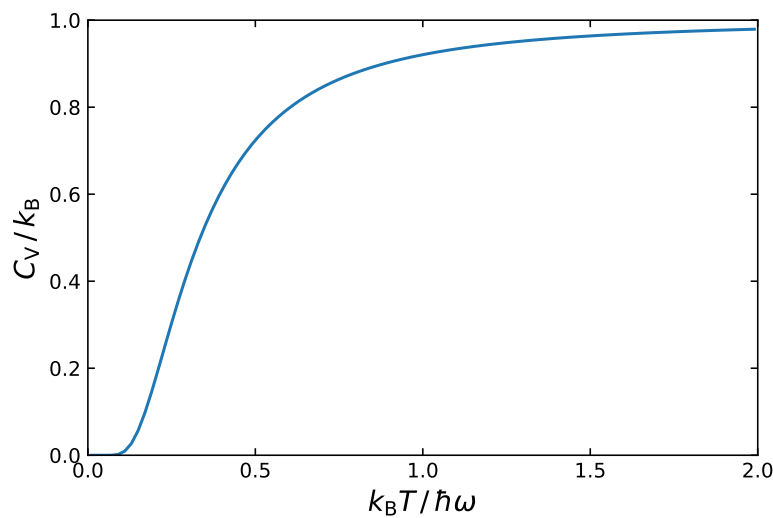


Figure 2: Heat capacity for a 1D quantum harmonic oscillator

4.7.1 The Einstein solid

The heat capacity for the 1D harmonic oscillator is given by

$$C_V = \left(\frac{\partial Q}{\partial T} \right)_V = \left(\frac{\partial U}{\partial T} \right)_V.$$

Substituting $\beta = 1/k_B T$ in the expression for the internal energy and differentiating gives

$$C_V = \frac{1}{(e^{\hbar\omega/k_B T} - 1)^2} e^{\hbar\omega/k_B T} \left(\frac{\hbar\omega}{k_B T^2} \right) \hbar\omega.$$

Defining $\theta_E = \hbar\omega/k_B$ gives

$$C_V = \frac{e^{\theta_E/T}}{(e^{\theta_E/T} - 1)^2} \left(\frac{\theta_E}{T} \right)^2 k_B.$$

The Einstein solid model treats each atom in the solid as a 3D quantum harmonic oscillator (equivalent to three 1D oscillators). All the oscillators have the same frequency ω . For N atoms, this results in a heat capacity of $3NC_V$. The parameter θ_E is referred to as the Einstein temperature and is a characteristic of the solid.

5 Systems with variable particle number (Lecture 7)

5.1 The grand canonical ensemble

So far we have considered closed systems. For systems that are completely isolated we used the microcanonical ensemble and for systems in thermal equilibrium with a heat bath we used the canonical ensemble. We shall now generalize further to consider systems in which the number of particles in the system is not fixed. We will follow the maximum entropy approach we used when deriving the Boltzmann distribution.

Consider an assembly of N_A identical systems. The systems are in contact with each other. But, as well as heat, the systems can now also exchange particles with each other so the number of particles in each system is not fixed. Each system has a set of microstates, but because both the number of particles and the energy are variable, we need to specify both for a microstate. Each microstate is therefore characterized by an energy ϵ_j and a particle number N_j . The whole assembly is isolated so the total energy and the total number of particles in the assembly are fixed. The systems are distinguishable, and so the statistical weight of a configuration of the assembly is again given by

$$\Omega_n = \frac{N_A!}{\prod_j n_j!},$$

where n_j is the number of systems of the assembly that are in the j^{th} microstate. To find the most likely distribution, we again maximize the entropy, subject to the constraints. The constraints are now

$$\sum_j n_j = N_A, \quad \sum_j n_j \epsilon_j = N_A U, \quad \sum_j n_j N_j = N_A N,$$

where ϵ_j and N_j are respectively the energy and the number of particles for the j microstate of a system, U is the mean energy per system and N is the mean number of particles per system.

For the Lagrangian we now have

$$\ln \Omega_n - \alpha \sum_j n_j - \beta \sum_j n_j \epsilon_j - \lambda \sum_j n_j N_j.$$

Maximizing the Lagrangian,

$$\frac{\partial \ln \Omega_n}{\partial n_j} - \alpha - \beta \epsilon_j - \lambda N_j = 0 \quad \text{for all } j,$$

solving which (using Stirling) gives

$$n_j = e^{-\alpha - \beta \epsilon_j - \lambda N_j}.$$

The probability that a system is in the j^{th} microstate with energy ϵ_j and particle number N_j is therefore

$$p_j = \frac{n_j}{N_A} = \frac{e^{-\alpha}}{N_A} e^{-\beta \epsilon_j - \lambda N_j}.$$

5.1.1 Grand partition function — significance of α

We can recognize $e^{-\alpha}/N_A$ as a normalization factor again,

$$\sum_j p_j = 1$$

$$\Rightarrow p_j = \frac{1}{\mathcal{Z}} e^{-\beta\epsilon_j - \lambda N_j}, \quad \text{where} \quad \mathcal{Z} = \sum_j e^{-\beta\epsilon_j - \lambda N_j}.$$

\mathcal{Z} is the *grand partition function*. It is analogous to the partition function in the canonical ensemble, but also includes particle number. The summation is again over all microstates of the system.

5.1.2 Temperature — significance of β

The entropy of the system is

$$\begin{aligned} S &= -k_B \sum_j p_j \ln p_j \\ &= -k_B \sum_j p_j (-\beta\epsilon_j - \lambda N_j - \ln \mathcal{Z}) \\ &= k_B \beta \sum_j p_j \epsilon_j + k_B \lambda \sum_j p_j N_j + k_B \ln \mathcal{Z} \sum_j p_j \\ &= k_B \beta U + k_B \lambda N + k_B \ln \mathcal{Z}, \end{aligned}$$

where U is the mean energy (the internal energy) of the system and N is the mean number of particles in the system.

For fixed V (so that the energy levels do not change) and fixed N ,

$$dS = k_B \beta dU \quad \Rightarrow \quad \left(\frac{\partial S}{\partial U} \right)_{N,V} = k_B \beta.$$

and we identify $\beta = 1/k_B T$, as before.

5.1.3 Chemical potential — significance of λ

The Lagrange multiplier λ was introduced as a constraint on the number of particles in the system. For V and U fixed,

$$dS = k_B \lambda dN \quad \Rightarrow \quad \left(\frac{\partial S}{\partial N} \right)_{U,V} = k_B \lambda.$$

In thermodynamics, for systems with variable particle number, the fundamental equation of thermodynamics becomes

$$dU = TdS - PdV + \mu dN,$$

where μ is the chemical potential. This gives

$$\left(\frac{\partial S}{\partial N} \right)_{U,V} = -\frac{\mu}{T}.$$

We can therefore relate the Lagrange constant λ to the chemical potential of thermodynamics: $\lambda = -\mu/k_B T$.

Rewriting in terms of the chemical potential, we now have for the probability that the system is in the j^{th} microstate

$$p_j = \frac{1}{\mathcal{Z}} e^{-(\epsilon_j - \mu N_j)/k_B T},$$

where the grand partition function is

$$\mathcal{Z} = \sum_j e^{-(\epsilon_j - \mu N_j)/k_B T}.$$

The chemical potential is a potential for particle number in the system. It is an energy per particle associated with adding particles to the system. For systems that can exchange particles, the chemical potential plays a similar role for particle number that electric potential does for charge. Particles will tend to flow from the higher chemical potential to the lower. Two systems that are in equilibrium must have the same chemical potential.

5.2 Grand potential

As with the partition function in the canonical ensemble, we shall see that the grand partition function can be related directly to a thermodynamic potential. Returning to the expression for the entropy of the system and substituting $\lambda = -\mu/k_B T$ and $\beta = 1/k_B T$, we have

$$S = \frac{1}{T} U - \frac{\mu}{T} N + k_B \ln \mathcal{Z},$$

from which we can see that

$$-k_B T \ln \mathcal{Z} = U - TS - \mu N.$$

Thus we have an expression relating the grand partition function to a thermodynamic potential, the *grand potential*,

$$\Phi = U - TS - \mu N.$$

The grand partition function is the key link to thermodynamics for systems described by the grand canonical ensemble,

$$\Phi = -k_B T \ln \mathcal{Z}.$$

Therefore if we know \mathcal{Z} , we can use it to calculate any thermodynamic variables via the grand potential. For example, using the fundamental equation of thermodynamics for a system,

$$dU = TdS - PdV + \mu dN \quad \Rightarrow \quad d\Phi = -SdT - PdV - Nd\mu.$$

Taking derivatives we obtain

$$S = - \left(\frac{\partial \Phi}{\partial T} \right)_{V, \mu}, \quad P = - \left(\frac{\partial \Phi}{\partial V} \right)_{T, \mu}, \quad N = - \left(\frac{\partial \Phi}{\partial \mu} \right)_{V, T}.$$

Appendix: Some mathematical tools

Math 1: Stirling's Formula

For $x \gg 1$,

$$\begin{aligned}\ln x! &= x \ln x - x + \frac{1}{2} \ln x + \frac{1}{2} \ln 2\pi + O(x^{-1}) \\ &\approx x \ln x - x\end{aligned}$$

We often require differentiation of $\ln x!$. Using Stirling's formula gives

$$\frac{d}{dx} \ln x! \approx \frac{d}{dx} (x \ln x - x) = \ln x$$

Math 2: Combinatorics

Permutations: The number of ways n distinct items can be ordered is

$$n!$$

Example: The number of anagrams of PHYSICAL is $8!$.

Permutations with repetition: The number of possible sequences of length n from a set of k items where each item can be chosen multiple times is

$$k^n$$

Example: The number of 4-letter words from a 26-letter alphabet is 26^4 .

Example: The number of microstates in a system of N distinguishable particles with m single-particle states. For each particle there is a choice of m states. The total number of microstates is then

$$\Omega = m^N$$

Combinations: The number of ways of choosing n items from a set of k items where the order does not matter is

$$\binom{n}{k} = \frac{k!}{(k-n)!n!} \quad (\text{the binomial coefficient})$$

Example: The number of ways of getting exactly n heads from k coin-tosses.

Permutations of multisets: The number of ways of choosing an ordered sequence of length n from a set of k different items, where each item must appear exactly m_i times is

$$\binom{n}{m_1, m_2, \dots, m_k} = \frac{n!}{m_1! m_2! \dots m_k!}$$

Example: The number of possible messages with length 22 where 'G' appears 5 times, 'C' 10 times, 'A' 3 times and 'T' 4 times, is $22!/(5! 10! 3! 4!)$

Example: The number of microstates in a configuration of N distinguishable particles specified by the occupancies $[n_0, n_1, \dots, n_k]$ is

$$\Omega_n = \frac{N!}{n_0! n_1! \dots n_k!}$$

Combinations with repetition: The number of ways of choosing n items from a set of k where each item can be chosen multiple times and the order does not matter is

$$\frac{(n+k-1)!}{n!(k-1)!}$$

Example: The number of microstates for a system with n identical bosons in k single-particle states. Specifying the state of every boson gives n choices of m states. The order does not matter since the particles are indistinguishable.

Math 3: The method of Lagrange multipliers

Consider a function $f(\mathbf{r})$ that is to be maximized subject to a constraint $g(\mathbf{r}) = c$, where c is a constant. If the gradients of the functions f and g are not parallel at a point \mathbf{r} then it is possible to find a vector \mathbf{v} such that

$$\mathbf{v} \cdot \nabla f(\mathbf{r}) \neq 0 \quad \text{and} \quad \mathbf{v} \cdot \nabla g(\mathbf{r}) = 0$$

Then moving in the direction of \mathbf{v} would change the value of f but would not change g which means that \mathbf{r} cannot be the position of a maximum. Therefore a necessary condition for a maximum in f subject to the constraint on g is that their gradients are parallel. If two vectors are parallel, they must be equal apart from a scalar multiplier. That is,

$$\nabla f(\mathbf{r}) = \lambda \nabla g(\mathbf{r}) \quad \text{for some multiplier } \lambda.$$

This generalizes to more than one constraint. Consider the case where we now have multiple constraints, represented by a set of constraint equations $g_i(\mathbf{r}) = c_i$. If it is possible to find a vector \mathbf{v} such that

$$\mathbf{v} \cdot \nabla f(\mathbf{r}) \neq 0 \quad \text{and} \quad \mathbf{v} \cdot \nabla g_i(\mathbf{r}) = 0 \quad \text{for every } g_i,$$

then by moving in the direction of \mathbf{v} it would be possible to change f , but not change any of the g_i , and so point \mathbf{r} cannot be a maximum. A necessary condition for the maximum therefore is that any vector \mathbf{v} that is perpendicular to all the constraints $\nabla g_i(\mathbf{r})$ must also be perpendicular to $\nabla f(\mathbf{r})$. This requires that $\nabla f(\mathbf{r})$ must lie in the space of directions defined by the constraint gradients, and can therefore be written in the form

$$\nabla f(\mathbf{r}) = \sum_i \lambda_i \nabla g_i(\mathbf{r}).$$

We can rewrite this condition as

$$\nabla \left[f(\mathbf{r}) - \sum_i \lambda_i g_i(\mathbf{r}) \right] = 0$$

where the function in brackets is called the Lagrangian, which for N constraints is

$$L(\mathbf{r}, \lambda_1, \lambda_2, \dots, \lambda_N) = f(\mathbf{r}) - \sum_{i=1}^N \lambda_i g_i(\mathbf{r})$$

and where the λ_i are the Lagrange multipliers. Thus we have transformed our original problem of solving a constrained maximization problem to one of an unconstrained extremization of the Lagrangian.

Note that although this is a necessary condition for the constrained maximum, it is not the complete solution. Solving this equation can find sets of solutions. It is then necessary to substitute in the values of the constraints (in this case the constants c_i) to obtain the correct solutions.