

Statistical Physics

Spring 2010

Dr Arttu Rajantie

January 6, 2010

Contents

1	Microstates and Macrostates	2
1.1	Microscopic Description	2
1.2	Macroscopic Description	3
1.3	Statistical Physics	3
2	Isolated Systems	4
2.1	Microcanonical Ensemble	4
2.2	Distinguishable and Indistinguishable Particles	5
2.3	Bosons and Fermions	7
2.4	Entropy and the Second Law	8
3	Systems with Variable Energy	12
3.1	Zeroth Law	12
3.2	Canonical ensemble	13
3.3	Connection with Thermodynamics	15
3.4	Single Free Particle	18
3.5	Gas of Distinguishable Particles	21
4	Variable Particle Number	22
4.1	Grand Canonical Ensemble	22
4.2	Bose-Einstein and Fermi-Dirac Distributions	25
5	Bosonic Gases	27
5.1	Black-body Radiation	27
5.2	Bose-Einstein Condensation	30
6	Fermionic Gases	34
6.1	Degenerate Fermion Gas	34
6.2	Non-Relativistic Degenerate Electron Gas	35
6.3	Ultrarelativistic Degenerate Electron Gas	36
6.4	White Dwarf Stars	36

A	Mathematical Results	39
A.1	Combinatorics	39
A.2	Stirling's Formula	39
A.3	Gaussian Integrals	40
A.4	Other Useful Integrals	40
A.5	Sums	41
A.5.1	Geometric Series	41
A.5.2	Multiple Summation	41
B	General Expression for Entropy	43
C	Notation	45
C.1	Indices	45
C.2	Quantities	45
C.3	Constants	45

1 Microstates and Macrostates

1.1 Microscopic Description

Last term in the Thermodynamics course, you learned about the physics of **macroscopic** systems, consisting of a **large number of particles** $N \gg 1$. It could be a drop of water, a cup of coffee, some gas in a piston or the whole Universe. You were considering the macroscopic properties of the system such as temperature, pressure etc. However, we all know that matter is really made of atoms, and at least in classical mechanics we could in principle go and measure the **position** \vec{x}_i and **velocity** of each of them \vec{v}_i . This would obviously be a much more detailed description of the state of the system, and if we assume that the atoms are pointlike, it would describe the state completely. In general, we call **the complete description** of the state of the system, containing all possible information about it, its **microstate**.

Within this **microscopic** description, we can find out exactly how the system behaves if we calculate the forces $\vec{F}_i(\vec{x}_1, \dots, \vec{x}_N)$ acting on all the particles, and then solve Newton's second law

$$m \frac{d^2 \vec{x}_i}{dt^2} = \vec{F}_i(\vec{x}_1, \dots, \vec{x}_N).$$

This gives us an **exact description of the behaviour of the system**. This approach is known as molecular dynamics, and in principle it is superior to the macroscopic thermodynamic description, which is only approximate. However, it is usually not possible in practice because of the large number of particles (one decilitre of water contains $N = 3 \times 10^{24}$ water molecules), because one needs to solve the same number of coupled differential equations.

Furthermore, nature is not classical. In **quantum mechanics**, the **microstate** of the system is specified by its **quantum wave function** $\psi(\vec{x}_1, \dots, \vec{x}_N)$. We would have to solve the Schrodinger equation for all the N particles simultaneously. Instead of N coupled ordinary differential equations, one has an N -dimensional partial differential equation, which is even more hopeless. The largest quantum molecular dynamics simulations have described only around 1000 atoms.

1.2 Macroscopic Description

Thermodynamics proposes a very different approach. Instead of the microstate, one is only considering the **macrostate** of the system. The macrostate is the approximate state of the system, specified by **macroscopic variables** such as **internal energy** U , **volume** V , **temperature** T , **pressure** P , **entropy** S , ... Which particular variables one uses depends on the system and the kind of questions one is interested in. The macrostates satisfy the general empirical **laws of thermodynamics**

- 0: If A is in thermal equilibrium with B , and B is in thermal equilibrium with C , then A is in thermal equilibrium with C .
- 1: Conservation of energy in an isolated system.
- 2: Heat cannot pass from a colder to a hotter body by itself.
- 3: All substances have the same entropy at absolute zero.

and certain other relations obtained from experiments, such as the **equation of state** ($PV = Nk_B T$ for an ideal gas).

This is an amazingly accurate theory of **macrostates** of systems in thermal equilibrium, but it raises some questions:

- (i) **Where did the laws of thermodynamics come from?**
- (ii) **What is the microscopic meaning of the thermodynamic variables, especially T and S ?**
- (iii) **Why does thermodynamics work so well?**
- (iv) **If we know what the system is made of, how do we derive the equation of state?**

These questions are important partly because if we cannot answer them we clearly do not really understand how nature works. However, they are also of concrete practical importance because being able to calculate the thermodynamic properties of the system allows us to make predictions about its behaviour in different conditions, such as very low temperatures, and to describe phenomena and systems that cannot be studied experimentally such as the early universe or the interior of a neutron star.

1.3 Statistical Physics

The **aim of statistical physics** is to answer these questions and thereby **provide a link between the microscopic and macroscopic descriptions**. We do that by using statistical methods.

It is clear that the system has many more microstates than macrostates. The air in a lecture theatre has (to a good approximation) the same temperature, pressure etc. throughout the lecture and therefore it stays in the same macrostate. However, the molecules that make up the air are constantly moving around in the room, and therefore the microstate is constantly changing. All these microstates correspond to the same macrostate.

There are therefore **a large number of microstates** corresponding to each **macrostate**. This collection of microstates is known as an **ensemble**. As we will see later, **different microstates** may occur with **different probabilities** in a given macrostate, and we have to take this into account. Mathematically, we define the ensemble as a probability distribution in the space of all possible microstates. It associates a probability p_α with each microstate α .

The evolution of the ensemble under the microscopic laws is usually much simpler than the evolution of the microstates. In particular, if the ensemble does not change with time at all, the system is said to be in an **equilibrium** state. Even then, the microstate of the system changes constantly according to the microscopic equations of motion.

The ensemble is therefore a statistical description of the macrostate of the system, and every macrostate corresponds to some particular ensemble. The key questions in statistical physics are

- (i) What is the correct ensemble to describe a given macrostate?
- (ii) How do we calculate the thermodynamic properties of the system from a given ensemble?

2 Isolated Systems

2.1 Microcanonical Ensemble

Some of the thermodynamic variables are easier to interpret microscopically than others. In particular, volume V and particle number N are quantities that exist and are well defined for a single microstate. The same is also true for **internal energy** U , which can be identified with the total microscopic energy E of the system, normalised in such a way that the minimum accessible energy is zero,

$$U = E - E_{\min}. \quad (2.1)$$

Therefore, we start by considering an **isolated system**, by which we mean a system with fixed U , N and V .

The **macrostate** of an isolated system is specified by the values of U , N , and V , and therefore the energy, particle number and volume of any microstate in the corresponding ensemble have to have these same values. In general, there is a large number of such microstates, and we have to choose what statistical weights we give to them. Since we do not have any more information about the system, we assume that

all possible microstates of an isolated system in equilibrium are equally probable.

This is the **fundamental postulate** of statistical mechanics. It is just a postulate, and it is not always true, but in practice it works very well. The ensemble defined in this way is known as the **microcanonical ensemble**. Mathematically, the microcanonical ensemble is defined as the probability distribution

$$p_{\alpha} = \begin{cases} p_0, & \text{if } E_{\alpha} = U, V_{\alpha} = V \text{ and } N_{\alpha} = N \\ 0 & \text{otherwise,} \end{cases} \quad (2.2)$$

where E_{α} , V_{α} and N_{α} are the energy, volume and particle number of microstate α , and p_0 is a constant.

As such, the fundamental postulate only makes sense if the number of microstates with given U , N and V is finite, but in classical mechanics it is usually not even countable. This problem is usually avoided in quantum mechanics, since observables are quantised.

As an example, let us consider a very simple case, a simple harmonic oscillator. Classically, its energy is

$$\epsilon = \frac{1}{2}mv^2 + \frac{1}{2}m\omega^2x^2.$$

The set of microstates that have a given energy ϵ is an ellipse, and being continuous, contains an infinite number of points. We would therefore have to choose what probability distribution we use on the ellipse.

M	Ω	microstates
0	1	(0, 0, 0)
1	3	(1, 0, 0), (0, 1, 0), (0, 0, 1)
2	6	(1, 1, 0), (1, 0, 1), (0, 1, 1), (2, 0, 0), (0, 2, 0), (0, 0, 2)
3	10	(1, 1, 1), (2, 1, 0), (2, 0, 1), (1, 2, 0), (0, 2, 1) (1, 0, 2), (0, 1, 2), (3, 0, 0), (0, 3, 0), (0, 0, 3)

 Table 1: Multiplicities for $N = 3$ distinguishable harmonic oscillators.

In **quantum mechanics**, we identify microstates with different orthogonal quantum states. In practice this means eigenstates of energy or some other observable. A simple harmonic oscillator has a discrete set of energy eigenstates

$$\epsilon_r = \hbar\omega \left(r + \frac{1}{2} \right),$$

where r is a non-negative integer. In a system that consists of a single harmonic oscillator, each one of these states is a different microstate. We are more interested systems with large numbers of particles (or oscillators), and we refer to the possible states of a single particle as **single-particle states** or often just “states” for brevity. It is very important to understand the difference between a microstate and a single-particle state.

In what follows, we will assume that the number of microstates in the microcanonical ensemble is finite. This number, which we denote by Ω , is called the **multiplicity** of the macrostate. It is the central statistical quantity in the study of isolated systems. Because the probabilities of all microstates have to add up to one, the probability p_0 in Eq. (2.2) is $p_0 = 1/\Omega$.

2.2 Distinguishable and Indistinguishable Particles

In counting the microstates, it makes a big difference whether the particles (or whatever the constituents of the system are) are **distinguishable**.

To see this, let us consider a set of N identical oscillators, which interact very weakly so that they can exchange energy but the energy levels of the oscillators are not affected. We first assume that they are positioned in a row, so that we can label by $i = 1, \dots, N$ according to their position. We denote the state of oscillator i by r_i . The microstate α of the whole system is specified by listing the states of all N oscillator, and therefore we can think of it as an N -component vector $\alpha = (r_1, \dots, r_N)$. Using Eq. (2.1), we define the internal energy as the total energy minus the zero-point energy,

$$U = E - E_{\min} = \sum_{i=1}^N (\epsilon_{r_i} - \epsilon_0) = \hbar\omega \sum_{i=1}^N r_i.$$

Consider now the microcanonical ensemble with energy $U = \hbar\omega M$. It contains all microstates in which the states r_i of the individual oscillators satisfy

$$\sum_{i=1}^N r_i = M.$$

As a concrete example, let us take $N = 3$. In that case, we can represent the microstate α of the system by a three-component vector $\alpha = (r_1, r_2, r_3)$. The multiplicities of macrostates with $M =$

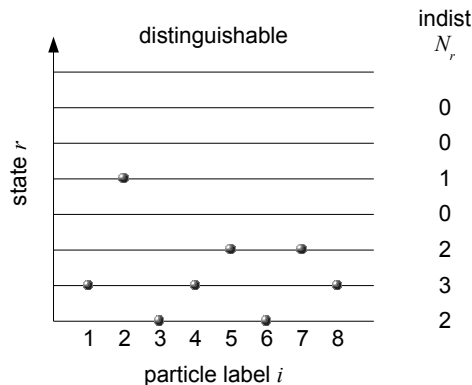


Figure 1: If the particles are indistinguishable, the microstates are fully described by the occupation numbers n_r , which give the number of particles in state r (right). If the particles are distinguishable, we would have to know the state of each particle (left).

0, 1, 2 and 3, as well as the actual microstates are listed in Table 1. In this case, since the oscillators are labelled by their position, we can distinguish between them, and we call them **distinguishable**. The state $(1, 0, 0)$ is therefore physically different from state $(0, 1, 0)$ and so on.

Consider now a system of N atoms in the same harmonic potential well around the same point in space. This is what is done in Bose-Einstein condensate experiments, which we will come back to later. In classical mechanics, each atom would still have its own identity, since we would be able to trace its motion and label the particles according to their positions at some reference time t_0 . For instance, it would be a well defined question to ask where the particle that was at position x_0 at time t_0 is at some later time t .

However, we cannot do this in quantum mechanics. The atoms do not have well-defined trajectories, and there is no sense in which we can label them or distinguish between them. The atoms are **indistinguishable**. If we swap the positions of two particles, we end up in the same physical state. This means that if we again represent the microstate α by integers r_i , different permutations of the set of integers r_i correspond to the exactly the same physical situation and should be counted as the one and the same microstate. Instead of a vector, in which the order of the components matters, we can represent the microstate by an unordered set consisting of the integers r_i . To emphasize the difference, we write it with curly brackets, $\alpha = \{r_1, r_2, r_3\}$. Since $\{1, 0, 0\} = \{0, 1, 0\} = \{0, 0, 1\}$, it is useful to choose that we always write them in, say, decreasing order. Then each state has a unique representation.

Another way to represent the microstate of a system is to tell how many particles are in each single-particle state r . This is known as the **occupation number** of state r and denoted by n_r . Since the particles are indistinguishable, it makes no sense to ask **which** particles are in each state. Therefore, the occupation numbers of all single-particle states specify the microstate completely. The set of occupation numbers forms an ordered list, $\alpha = [n_0, n_1, n_2, \dots]$, and this is often the most convenient representation for the microstate of a system of indistinguishable particles. For clarity, we use square brackets when listing occupation numbers. For an illustration, see Fig. 1.

As we will see, distinguishable and indistinguishable particles lead to very different macroscopic

dynamics. As always in physics, it is ultimately experiments rather than the theoretical reasoning given above that tell us that identical particles have to be treated as indistinguishable. It is remarkable that relatively simple macroscopic experiments can reveal such a fundamental property of nature.

2.3 Bosons and Fermions

Besides indistinguishability of particles, quantum mechanics leads to a further complication, which has equally profound implications for macroscopic physics. Consider the quantum wave function $\psi(x_1, x_2)$ of two identical particles. As usual, the probability to find particles at positions x_1 and x_2 is given by $|\psi(x_1, x_2)|^2$. Since swapping the positions of the two particles does not change the state, we must have

$$|\psi(x_1, x_2)|^2 = |\psi(x_2, x_1)|^2. \quad (2.3)$$

This requires

$$\psi(x_2, x_1) = e^{i\theta} \psi(x_1, x_2), \quad (2.4)$$

where θ is some real number. If we repeat this, we find

$$\psi(x_1, x_2) = \left(e^{i\theta}\right)^2 \psi(x_1, x_2), \quad (2.5)$$

and therefore $e^{i\theta} = \pm 1$. This means that

$$\psi(x_2, x_1) = \pm \psi(x_1, x_2), \quad (2.6)$$

where the sign is an intrinsic property of the given particle species.

Let us now assume that the two particles are in states $\psi_A(x)$ and $\psi_B(x)$. The general two-particle wave function is a linear combination

$$\psi(x_1, x_2) = c_1 \psi_A(x_1) \psi_B(x_2) + c_2 \psi_B(x_1) \psi_A(x_2), \quad (2.7)$$

in which c_1 and c_2 are constants. The first term corresponds to particle 1 being in state A and particle 2 in state B, and the second term to particle 1 being in state B and particle 2 in state A. Using Eq. (2.6), we find $c_1 = \pm c_2$, and consequently

$$\psi(x_1, x_2) = c_1 [\psi_A(x_1) \psi_B(x_2) \pm \psi_B(x_1) \psi_A(x_2)]. \quad (2.8)$$

Finally, let us assume that the states ψ_A and ψ_B are the same, which means that we have two particles in the same state. For the + sign, we find

$$\psi(x_1, x_2) = 2c_1 \psi_A(x_1) \psi_A(x_2), \quad (2.9)$$

which is perfectly fine. However, for the – sign,

$$\psi(x_1, x_2) = c_1 [\psi_A(x_1) \psi_A(x_2) - \psi_A(x_1) \psi_A(x_2)] = 0, \quad (2.10)$$

which means that the state has zero amplitude and does not exist.

Thus, it follows from quantum mechanics that there are two types of particles, and their counting is different:

- **Bosons** correspond to the + sign. The number of bosons in a given state is not constrained.

Bosons		
M	Ω	microstates
0	1	$\{0, 0, 0\} = [3, 0, 0, \dots]$
1	1	$\{1, 0, 0\} = [2, 1, 0, \dots]$
2	2	$\{1, 1, 0\} = [1, 2, 0, \dots]$, $\{2, 0, 0\} = [2, 0, 1, \dots]$
3	3	$\{1, 1, 1\} = [0, 3, 0, \dots]$, $\{2, 1, 0\} = [1, 1, 1, \dots]$, $\{3, 0, 0\} = [2, 0, 1, \dots]$
Fermions		
M	Ω	microstates
0	0	-
1	0	-
2	0	-
3	1	$\{2, 1, 0\} = [1, 1, 1, 0, \dots]$
4	1	$\{3, 1, 0\} = [1, 1, 0, 1, \dots]$
5	2	$\{3, 2, 0\} = [1, 0, 1, 1, \dots]$, $\{4, 1, 0\} = [1, 1, 0, 1, \dots]$

Table 2: Multiplicities for $N = 3$ indistinguishable harmonic oscillators.

- **Fermions** correspond to the $-$ sign. They obey the **Pauli exclusion principle**, which means that two particles cannot be in the same state.

It can be proven that this statistical nature of particles is fundamentally related to their spin. Particles with integer spin are bosons, and particles with half-integer spin such as $1/2$ or $3/2$ are fermions. Examples of fermions are electrons, protons and neutrons. When a particle is made of several elementary particles, the spins are added together modulo an integer, and therefore a particle that contains an odd number of fermions is a fermion. For example, ^3He atoms, which are made of two protons, one neutron and two electrons, are fermions. Examples of bosons are photons, hydrogen atoms (one proton and one electron) and ^4He atoms (two protons, two neutrons and two electrons).

When counting the multiplicities of macrostates, one has to keep in mind whether the particles in question are bosons or fermions, since in the latter case all particles have to be in different single-particle states. Table 2 shows the multiplicities of the macrostates on a three-particle system with different M and the corresponding microstates. The microstates are shown both by listing the states $\{r_1, r_2, r_3\}$ and by listing the occupation numbers $[n_0, n_1, \dots]$. A general expression for the multiplicity is not known even for this simple case, which illustrates why microcanonical ensemble is not well suited for practical calculations.

2.4 Entropy and the Second Law

To understand the macroscopic significance of Ω , let us consider an example that is familiar from the Thermodynamics course. We have an isolated container of volume V_1 that is split into two parts by a wall (see Fig. 2). The part on the left has volume V_0 and is filled with gas, whereas the part on the right is empty.

If the wall is removed, the gas fills the whole container. This process is known as adiabatic free expansion. In the Thermodynamics course, you showed using the ideal gas equation of state that in

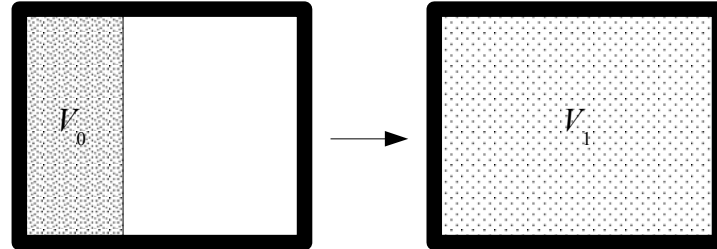


Figure 2: Adiabatic free expansion.

this process the entropy S of the system grows by

$$S_1 - S_0 = \Delta S = Nk_B \ln \frac{V_1}{V_0}, \quad (2.11)$$

where S_1 and S_0 are the entropies of the final and initial state, respectively.

One formulation of the second law of thermodynamics is that the entropy of an isolated system cannot decrease. The growth of entropy therefore means that adiabatic free expansion is an irreversible process. It is not possible for the system to go back from its final to its initial state.

That was how the process of adiabatic free expansion looks like in thermodynamics. Let us now think about it from the point of view of statistical physics. In the initial state, we have some number N of gas particles in volume V_0 with total energy U . This macrostate has some multiplicity, which we denote by Ω_0 . At this stage, we do not need to know how to compute it. In the final state, the same N particles are in volume V_1 with the same energy U , since the energy is conserved in an isolated system. This macrostate has a different multiplicity Ω_1 .

Since the Ω_1 microstates that correspond to the final state include all microstates with total energy U and N particles anywhere in the container, it also includes the Ω_0 microstates that correspond to the initial state. They are the microstates in which all particles happen to be in the original volume V_0 . Therefore, the fundamental postulate implies that the probability that the process is reversed, i.e., all the particles are in the original volume V_0 , is

$$p_N = \frac{\Omega_0}{\Omega_1}. \quad (2.12)$$

On the other hand, we can compute this probability directly. In the final state, a given particle can equally well be anywhere in the system. In particular, the probability p_1 that it is in the original volume V_0 is there given simply by the ratio of the volumes,

$$p_1 = \frac{V_0}{V_1}. \quad (2.13)$$

The probability p_N that **all** the particles are in the original volume V_0 is then the combined probability of the N individual ones,

$$p_N = \left(\frac{V_0}{V_1} \right)^N. \quad (2.14)$$

If N is large, then this probability is tiny. This is the microscopic reason why the process is **irreversible**. The system never returns into its original state in practice, although it is possible in principle.

Combining Eqs. (2.12) and (2.14), we find

$$\frac{\Omega_0}{\Omega_1} = \left(\frac{V_0}{V_1}\right)^N. \quad (2.15)$$

The logarithm of this is

$$N \ln \frac{V_0}{V_1} = \ln \frac{\Omega_0}{\Omega_1}. \quad (2.16)$$

Using this result, we can write Eq. (2.11) as

$$S_1 - S_0 = k_B \ln \frac{\Omega_1}{\Omega_0} = k_B \ln \Omega_1 - k_B \ln \Omega_0. \quad (2.17)$$

This suggests we identify

$$S = k_B \ln \Omega. \quad (2.18)$$

This equation was discovered by Ludwig Boltzmann, and the quantity S defined by the equation is the **Boltzmann entropy**. Note that in thermodynamics, the definition of entropy involves the concept of temperature, which itself is only properly defined for perfect gases. In contrast, Eq. (2.18) can be applied to any system, at least in principle, and is therefore a more general definition of entropy.

Equation (2.18) is one of the most important equations in statistical physics, since it provides a link between the macroscopic (left hand side) and microscopic (right hand side) descriptions of the system. When we know entropy, we can use it to compute other thermodynamic variables using the relations we have learned in thermodynamics. In particular, we use the relation

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_{V,N}, \quad (2.19)$$

which follows directly from the fundamental equation of thermodynamics. This gives a **universal and absolute** definition for the **temperature** of any system.

We can carry out a few checks to make sure that the quantity defined by Eq. (2.18) really behaves the way we would expect entropy to behave. First, the ground state of a quantum mechanical system is (usually) unique. Therefore, if the system is in its ground state, as we would expect it to be at zero temperature, the multiplicity is $\Omega = 1$. According to Eq. (2.18), this implies $S = 0$, in agreement with the third law of thermodynamics. Second, the quantity defined by Eq. (2.18) is extensive, as one can see by considering two systems A and B , whose multiplicities are Ω_A and Ω_B . If the systems do not interact or interact only weakly, their microstates are independent of each other, and therefore the total number of microstates is

$$\Omega_{\text{tot}} = \Omega_A \Omega_B. \quad (2.20)$$

Eq. (2.18) then implies

$$S_{\text{tot}} = S_A + S_B. \quad (2.21)$$

As an example, let us calculate the entropy of a system of N **distinguishable** spin-1/2 particles in magnetic field. Quantum mechanically, such a spin can be in two different states, either pointing in the direction of the magnetic field or opposite to it. We refer to these states as “down” and “up”. The

down state has a lower energy, and we choose it to be zero, and we label the energy of the up state by ϵ_0 . The energy of a single spin i is therefore $\epsilon_i = \epsilon_0 r_i$, where

$$r_i = \begin{cases} 1 & \text{for the up state and} \\ 0 & \text{for the down state.} \end{cases} \quad (2.22)$$

The total internal energy is therefore

$$U = \epsilon_0 \sum_{i=1}^N r_i. \quad (2.23)$$

The microcanonical ensemble corresponds to a fixed value of U , and therefore to a fixed value of

$$M \equiv \sum_{i=1}^N r_i. \quad (2.24)$$

To calculate the multiplicity of the macrostate corresponding to a given M , we have to count all the possible ways the N spin can be arranged so that M of them point up. Let us first look at a few of the lowest possible values of M :

$M = 0$: The only microstate with $M = 0$ is the one with all spins down. Therefore $\Omega(0) = 1$.

$M = 1$: Start from $M = 0$ and flip one of the spins. There are N spins to choose from, so $\Omega(1) = N$.

$M = 2$: Flip one of the $(N - 1)$ spins that are still pointing down. There are $N(N - 1)$ possible ways to flip the two spins, but the final microstate is the same if you change the order in which you flipped them, so $\Omega(2) = N(N - 1)/2$.

...

The general expression for the multiplicity is given by the number of different ways M items can be chosen from the total of N if the order does not matter (see Classwork 1),

$$\Omega(M) = \frac{N!}{M!(N - M)!}. \quad (2.25)$$

The entropy of this system is therefore

$$S = k_B \ln \Omega = k_B (\ln N! - \ln M! - \ln(N - M)!). \quad (2.26)$$

When N , M and $N - M$ are all large, we can use Stirling's formula (see Appendix A)

$$\ln N! \approx N \ln N - N, \quad (2.27)$$

to simplify this into

$$S \approx k_B \left(M \ln \frac{N}{M} + (N - M) \ln \frac{N}{N - M} \right). \quad (2.28)$$

Some consequences of this result are explored in Problem Sheet 1.

Finally, let us discuss the implications of Boltzmann's definition (2.18) of entropy. One way to phrase the second law of thermodynamics is that the entropy of an isolated system cannot decrease. A process in which entropy grows is therefore irreversible. This defines an "arrow of time" since such

processes can only take place in one direction, even though the microscopic laws of nature are time reversal symmetric.

The identification of entropy with probability explains this. When the system evolves, it will generally pass through all possible microstates with the same U and N . (This is known as the **ergodic hypothesis**.) Unless we actually track the evolution of its microstates, the best we can say after a while is that it can equally well be in any microstate. This is the content of the fundamental postulate of statistical physics.

If we start in a macrostate that has a lower entropy than the equilibrium state, Eq. (2.18) says that the multiplicity of the initial state is smaller than the total number of possible microstates. After a while, the probabilities of all microstates have become equal, and the system is in the equilibrium state, which has maximum entropy S . Therefore the entropy always grows.

This irreversibility does not mean that the laws of physics distinguish between past and future. If we could actually reverse the velocities of all atoms at a later time, the system would evolve back into its original state. However, this would require dealing with the microstate of the system, and the second law only refers to the evolution of the macrostate. Likewise, if we wait long enough the system will eventually come pass through its original state, but we would have to wait an extremely long time.

Instead, what distinguishes between past and future is the way we have set the system up. If the system starts in a state with low entropy, we must have prepared it so. For instance, we have filled only part of the container with gas. This is an highly atypical state for the system and would not have occurred without our interference. Once we have done that, it is only natural that at a later time, we find the system to be in a more typical state.

The origin of the arrow of time, i.e., why all natural thermodynamics processes take place only in one direction, is therefore that the entropy of the Universe is relatively low, and therefore it is much more likely that it grows than that it decreases.

3 Systems with Variable Energy

3.1 Zeroth Law

As a warm-up exercise, consider an isolated container separated into two parts by a diathermal wall, meaning that it allows the two sides to exchange energy in the form of heat. The wall cannot move, and particles cannot penetrate it, so the volumes and particle numbers of the two sides remain fixed. Let us call the left hand side system 1 and the right hand side system 2.

Let us now consider a macrostate in which the energies of systems 1 and 2 are U_1 and U_2 . The total energy is the sum of these two

$$U = U_1 + U_2, \quad (3.1)$$

and since the whole system is isolated, it is fixed. If we assume that the two systems interact very weakly, their microstates are independent, and therefore the total multiplicity is the the product of the multiplicities of the two systems. Since the multiplicity of the system 1 depends only on U_1 , and the multiplicity of system 2 only on U_2 , we have

$$\Omega_{\text{tot}}(U_1, U_2) = \Omega_1(U_1)\Omega_2(U_2). \quad (3.2)$$

According to Eq. (2.18), we have

$$S_{\text{tot}}(U_1, U_2) = S_1(U_1) + S_2(U_2), \quad (3.3)$$

where $S_i(U_i)$ is the entropy of system i at energy U_i .

In equilibrium, entropy of the whole system is maximised, meaning that

$$\frac{\partial S_{\text{tot}}}{\partial U_1} = 0. \quad (3.4)$$

Using Eq. (3.3), we can write this as

$$\frac{\partial S_{\text{tot}}}{\partial U_1} = \frac{\partial S_1}{\partial U_1} + \frac{dU_2}{dU_1} \frac{\partial S_2}{\partial U_2} = \frac{\partial S_1}{\partial U_1} - \frac{\partial S_2}{\partial U_2}, \quad (3.5)$$

where we have used $U_2 = U - U_1$. Therefore, Eq. (3.4) becomes

$$\frac{\partial S_1}{\partial U_1} = \frac{\partial S_2}{\partial U_2}. \quad (3.6)$$

It then follows directly from the definition of temperature (2.19) that the temperatures of the two sides are equal

$$\frac{1}{T_1} = \frac{1}{T_2} \quad \Leftrightarrow \quad T_1 = T_2. \quad (3.7)$$

This means that two systems that are in thermal equilibrium with each other have the **same temperature**, which is equivalent with the zeroth law of thermodynamics. If you now consider a third system in equilibrium with system 3, then $T_3 = T_2$, and consequently $T_3 = T_1$, which means that systems 1 and 3 are in thermal equilibrium.

3.2 Canonical ensemble

Let us now make system 2 much larger than system 1. Its heat capacity, being an extensive quantity, becomes very high, and therefore its temperature remains practically constant $T_2 = T = \text{const}$. System 2 becomes a heat bath. Furthermore, we have

$$U_1 \ll U_2, \quad S_1 \ll S_2. \quad (3.8)$$

The probability that the energies of systems 1 and 2 are U_1 and U_2 is proportional to the multiplicity of such a macrostate,

$$p(U_1, U_2) \propto \Omega(U_1, U_2) = e^{S(U_1, U_2)/k_B} = e^{S_1(U_1)/k_B} e^{S_2(U_2)/k_B} = \Omega_1(U_1) e^{S_2(U_2)/k_B}, \quad (3.9)$$

where we used Boltzmann's formula (2.18) first for the whole system and then again for system 1.

It is now convenient to Taylor expand

$$S_2(U) = S_2(U_2 + U_1) = S_2(U_2) + \frac{\partial S_2}{\partial U_2} U_1 + \frac{1}{2} \frac{\partial^2 S_2}{\partial U_2^2} U_1^2 + \dots \quad (3.10)$$

By the definition (2.19) of temperature, the coefficient of the second term is

$$\frac{\partial S_2}{\partial U_2} = \frac{1}{T_2} = \frac{1}{T}. \quad (3.11)$$

Similarly, the coefficient of the third term is

$$\frac{1}{2} \frac{\partial^2 S_2}{\partial U_2^2} = \frac{1}{2} \frac{\partial}{\partial U_2} \frac{1}{T}, \quad (3.12)$$

which vanishes because the temperature T was assumed to be constant. Another way to see this would be to write $U_2 = u_2 N_2$, where u_2 is the energy per particle. Then we have

$$\frac{1}{2} \frac{\partial^2 S_2}{\partial U_2^2} = \frac{1}{2N_2} \frac{\partial(1/T)}{\partial u_2} \propto \frac{1}{N_2} \xrightarrow{N_2 \rightarrow \infty} 0, \quad (3.13)$$

where we have used the fact that T and u_2 are intensive quantities and therefore independent of N_2 .

Either way, this means that higher-order terms in Eq. (3.10) vanish, and we have

$$S_2(U) = S_2(U_2) + \frac{U_1}{T} \quad \Rightarrow \quad S_2(U_2) = S_2(U) - \frac{U_1}{T}. \quad (3.14)$$

Substituting this into Eq. (3.9), we find

$$p(U_1, U_2) \propto \Omega_1(U_1) e^{S_2(U)/k_B} e^{-U_1/k_B T}. \quad (3.15)$$

Note that since U is fixed the factor $\exp(S_2(U)/k_B)$ is just a constant. The expression depends only on U_1 , so we can drop the argument U_2 . We can therefore write

$$p(U_1) = \frac{1}{Z} \Omega_1(U_1) e^{-U_1/k_B T}, \quad (3.16)$$

where Z is a constant normalisation factor. This expression gives the probability that system 1 has energy U_1 . Since the expression is independent of details of system 2, the subscript 1 has become unnecessary, and we can conclude that in general, the probability that a system that is in thermal equilibrium with a heat bath at temperature T has energy U is

$$p(U) = \frac{1}{Z} \Omega(U) e^{-U/k_B T}. \quad (3.17)$$

Consider now a **given microstate** α that has energy E_α . The probability that the system is in this microstate is given by $p[E_\alpha]$ divided by the number of microstates with the same energy, which is equal to $\Omega[E_\alpha]$. This means

$$p_\alpha = \frac{1}{Z} e^{-E_\alpha/k_B T}. \quad (3.18)$$

This probability distribution is known as the **Boltzmann distribution**.

The normalisation factor Z can be computed from the condition that the sum of the probabilities of all possible microstates has to be one,

$$1 = \sum_{\alpha} p_\alpha = \frac{1}{Z} \sum_{\alpha} e^{-E_\alpha/k_B T}, \quad (3.19)$$

where the sum is over all microstates α . This implies

$$Z = \sum_{\alpha} e^{-E_\alpha/k_B T}. \quad (3.20)$$

Even though Z was introduced simply as a normalisation factor, we will see that it plays a crucial role in connecting statistical physics with thermodynamics. It is generally known as the **partition function**.¹

¹In textbooks, the partition function is sometimes written as a sum over *energy levels* rather than *microstates*. Then one has to include the factor $\Omega(U)$ in Eq. (3.17) to keep track of how many microstates have the same energy.

The probabilities in Eq. (3.18) define the **canonical ensemble**. In summary, the mathematical definition of the canonical ensemble is

$$p_\alpha = \begin{cases} \frac{1}{Z} e^{-E_\alpha/k_B T} & \text{if } V_\alpha = V \text{ and } N_\alpha = N, \\ 0 & \text{otherwise.} \end{cases} \quad (3.21)$$

Physically, the canonical ensemble models a system with fixed N and V in thermal equilibrium with a heat bath at temperature T . Note that the “system” in question can be even a single particle, in which case, we will denote its microstates by r and the corresponding energies by ϵ_r , so that

$$p_r = \frac{1}{Z} e^{-\epsilon_r/k_B T}. \quad (3.22)$$

Finally, it is often more natural to use the quantity

$$\beta = \frac{1}{k_B T} \quad (3.23)$$

instead of T . In terms of β , the probability of a given microstate α is

$$p_\alpha = \frac{1}{Z} e^{-\beta E_\alpha}. \quad (3.24)$$

3.3 Connection with Thermodynamics

Let us now try to use the canonical ensemble to calculate the thermodynamic properties of the system. The easiest quantity to compute is the internal energy U , since it corresponds simply to the **mean energy** \bar{E} of the probability distribution (3.18). Its value is therefore given by

$$U = \bar{E} = \sum_\alpha E_\alpha p_\alpha = \frac{1}{Z} \sum_\alpha E_\alpha e^{-\beta E_\alpha}. \quad (3.25)$$

Let us now note that if we differentiate Eq. (3.20) with respect to β , it brings down a factor $-E_\alpha$ and therefore we find

$$\frac{\partial Z}{\partial \beta} = \sum_\alpha (-E_\alpha) e^{-\beta E_\alpha} = - \sum_\alpha E_\alpha e^{-\beta E_\alpha}. \quad (3.26)$$

Comparing with Eq. (3.25), we can see that the two equations have exactly the same sum, and we can therefore write

$$\bar{E} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta}. \quad (3.27)$$

Writing this in terms of T instead of β , we have

$$\bar{E} = -\frac{dT}{d\beta} \frac{\partial \ln Z}{\partial T} = k_B T^2 \frac{\partial \ln Z}{\partial T}. \quad (3.28)$$

It is also straightforward to compute the **standard deviation** ΔE of energy in the canonical distribution, defined as

$$(\Delta E)^2 \equiv \overline{E^2} - \bar{E}^2. \quad (3.29)$$

First, we observe that in analogy with Eq. (3.26), we have

$$\frac{\partial^2 Z}{\partial \beta^2} = \sum_\alpha (E_\alpha)^2 e^{-\beta E_\alpha} = Z \overline{E^2}. \quad (3.30)$$

Taking the second derivative of the logarithm gives

$$\frac{\partial^2 \ln Z}{\partial \beta^2} = \frac{\partial}{\partial \beta} \left(\frac{1}{Z} \frac{\partial Z}{\partial \beta} \right) = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2} - \frac{1}{Z^2} \left(\frac{\partial Z}{\partial \beta} \right)^2 = \overline{E^2} - \overline{E}^2, \quad (3.31)$$

where in the last step we used Eqs. (3.27) and (3.30). Thus, we have

$$(\Delta E)^2 = \frac{\partial^2 \ln Z}{\partial \beta^2}. \quad (3.32)$$

Using Eq. (3.27) we can also write this as

$$(\Delta E)^2 = -\frac{\partial \overline{E}}{\partial \beta} = k_B T^2 C, \quad (3.33)$$

where $C = \partial U / \partial T$ is the heat capacity of the system. This implies that the relative fluctuation of energy is

$$\frac{\Delta E}{U} = \frac{\sqrt{k_B T^2 C}}{U}. \quad (3.34)$$

Both the heat capacity C and the internal energy U are extensive quantities, i.e., proportional to N . Therefore, the relative fluctuation depends on N as

$$\frac{\Delta E}{U} \propto \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}. \quad (3.35)$$

In the **thermodynamic limit** ($N \gg 1$) the relative fluctuation is therefore tiny. This means that the canonical ensemble is in fact, practically equivalent to the microcanonical one, in which the value of energy is precisely fixed.

One can compute the mean value of any other microscopic quantity X in the the same way, as

$$\overline{X} = \sum_{\alpha} p_{\alpha} X_{\alpha}. \quad (3.36)$$

This is, however, not really enough. To be able to relate our microscopic description to thermodynamics, we must be able to compute values of macroscopic thermodynamic variables such as S and P . To do that, we need the relevant thermodynamic potential, and we know from thermodynamics that for systems with variable energy that is the **free energy**² F , defined as

$$F = U - TS. \quad (3.37)$$

Once we know the free energy, we can calculate any thermodynamic variables. For example,

$$S = -\left(\frac{\partial F}{\partial T} \right)_V, \quad P = -\left(\frac{\partial F}{\partial V} \right)_T. \quad (3.38)$$

From the Boltzmann definition (2.18), one can derive a general expression (see Appendix B),

$$S = -k_B \sum_{\alpha} p_{\alpha} \ln p_{\alpha}, \quad (3.39)$$

²This quantity is sometimes also called the Helmholtz function.

which is valid for any ensemble, and we use this to calculate F . For the canonical ensemble (3.24), we have

$$\ln p_\alpha = -\beta E_\alpha - \ln Z \quad (3.40)$$

and therefore

$$S = k_B \left(\sum_\alpha p_\alpha \beta E_\alpha + \ln Z \right). \quad (3.41)$$

We note that the sum over microstates α gives the mean energy $\bar{E} = U$, so we have

$$S = k_B (\beta U + \ln Z) = \frac{U}{T} + k_B \ln Z. \quad (3.42)$$

Substituting this into Eq. (3.37) gives an explicit microscopic expression for the free energy,

$$F = -k_B T \ln Z. \quad (3.43)$$

This equation connects the thermodynamic description of the system (left hand side) with the microscopic description (right hand side). It is therefore analogous to the Boltzmann definition of entropy (2.18). Significantly, the quantity Z , which was introduced merely as a normalisation factor in Eq. (3.16), has turned out to play a key role in this connection.

As an example, let us calculate the internal energy U from Eq. (3.43). We use Eqs. (3.37) and (3.38) to write it as

$$U = F + TS = F - T \left(\frac{\partial F}{\partial T} \right)_V, \quad (3.44)$$

and substituting Eq. (3.43), we find

$$\begin{aligned} U &= -k_B T \ln Z + k_B T \left(\frac{\partial}{\partial T} T \ln Z \right)_V = -k_B T \ln Z + k_B T \ln Z + k_B T^2 \left(\frac{\partial}{\partial T} \ln Z \right)_V \\ &= k_B T^2 \left(\frac{\partial}{\partial T} \ln Z \right)_V, \end{aligned} \quad (3.45)$$

which agrees with the expression (3.28) we derived earlier directly from the probabilities p_α .

As a more concrete example, let us again consider a system of N distinguishable spin 1/2 particles. When we described it using the microcanonical ensemble, we had to calculate the multiplicities $\Omega(M)$. In this simple case, it was possible, but it is almost always tedious and often impossible. When we model the system using the canonical ensemble, we avoid this step altogether.

The partition function Z is given by Eq. (3.20), where the sum goes over all possible microstates of the system. Each of the spins can be in two possible states, $r_i = 0$ and $r_i = 1$, so the partition function is

$$Z = \sum_{r_1=0}^1 \cdots \sum_{r_N=0}^1 \exp \left(-\beta \epsilon_0 \sum_{i=1}^N r_i \right), \quad (3.46)$$

where I have used the notation \cdots which generally represents repeating the same expression. This sum factorises (see Appendix A) into

$$Z = \left(\sum_{r_1=0}^1 e^{-\beta \epsilon_0 r_1} \right) \cdots \left(\sum_{r_N=0}^1 e^{-\beta \epsilon_0 r_N} \right) = \left(1 + e^{-\beta \epsilon_0} \right)^N. \quad (3.47)$$

Now, we can use Eq. (3.43) to calculate the thermodynamic properties of the system. In contrast with the calculation done in Section 2.4, the whole calculation only involved mechanical manipulation of sums.

3.4 Single Free Particle

The canonical ensemble can be used to describe even a system consisting of a single particle. The ensemble is then just a probability distribution for the different states of the particle, which we will call **single-particle states** and label by r . If we denote the energy of state r by ϵ_r , the partition function is

$$Z_1 = \sum_r e^{-\beta\epsilon_r}, \quad (3.48)$$

where the subscript 1 indicates that this is the partition function of a single particle.

The single-particle states of free particles are labelled by their momentum \vec{p} and also by any internal degrees of freedom such as angular momentum, and we need to sum over all of them,

$$\sum_r = \sum_{\text{int}} \sum_{\vec{p}}. \quad (3.49)$$

The spectrum is continuous spectrum because the momentum p is not quantised. This means that the sum over \vec{p} should actually be an integral. However, when we replace the sum by an integral, we have to be careful to give each state the correct weight. This means that the sum over states corresponds to some **density of states** $f(\epsilon)$ defined by³

$$\sum_r \rightarrow \int d\epsilon f(\epsilon), \quad (3.50)$$

and we need to find $f(\epsilon)$. The concrete meaning of $f(\epsilon)$ is that the integral

$$\int_{\epsilon_0}^{\epsilon_1} d\epsilon f(\epsilon) \quad (3.51)$$

gives the number of states with energies between ϵ_0 and ϵ_1 .

In practice, the internal states and translational states are usually independent in the sense that energy of the single particle state is a sum of the two contributions

$$\epsilon_r = \epsilon_{\text{int}} + \epsilon_{\text{tr}}(\vec{p}).$$

In that case the partition function (3.48) factorises,

$$Z_1 = \left(\sum_{\text{int}} e^{-\beta\epsilon_{\text{int}}} \right) \left(\sum_{\vec{p}} e^{-\beta\epsilon_{\text{tr}}(\vec{p})} \right) \equiv Z_{\text{int}} Z_{\text{tr}}.$$

We will mostly consider the case in which the internal states are **degenerate**, which means that they have the same energy. Then $\epsilon_{\text{int}} = 0$, and Z_{int} is simply the number of different internal states, which is known as the **degeneracy** of the momentum state and is often denoted by g . If we now define the **density of translational states** $f_{\text{tr}}(\epsilon)$ by

$$\sum_{\vec{p}} \rightarrow \int d\epsilon f_{\text{tr}}(\epsilon),$$

³Sturge and Trevena denote the density of states by g .

we can see that the full density of states is

$$f(\epsilon) = g f_{\text{tr}}(\epsilon). \quad (3.52)$$

To calculate $f_{\text{tr}}(\epsilon)$, we imagine that our particle is in a finite box of size $V = L \times L \times L$. At the end of the calculation, we will take the box size to infinity, $L \rightarrow \infty$. In a finite box, the spectrum becomes quantised. The states of the particle correspond to standing waves that vanish at the walls. In one dimension, the wave functions are of the form

$$\psi(x) = A \sin \frac{2\pi x}{\lambda}, \quad (3.53)$$

where the wavelength λ is such that $\psi(L) = 0$. This gives the quantisation condition

$$\lambda_n = \frac{2L}{n}, \quad n \in \mathbb{Z}_+. \quad (3.54)$$

The momentum of the particle is then also quantised,

$$p_n = \frac{h}{\lambda_n} = \frac{hn}{2L}. \quad (3.55)$$

In three dimensions, the states are labelled by three positive integers (n_x, n_y, n_z) which we can think of forming a vector \vec{n} . The momentum of the particle in this state is

$$\vec{p}_{\vec{n}} = \frac{h\vec{n}}{2L} \equiv \Delta p \vec{n}, \quad (3.56)$$

where we have defined the **momentum spacing** $\Delta p = h/2L$. The possible states correspond to a cubic grid of points in the momentum space, with spacing Δp . When $L \rightarrow \infty$, the spacing goes to zero, $\Delta p \rightarrow 0$.

The (translational) energy of state \vec{n} is

$$\epsilon(\vec{p}_{\vec{n}}) = \sqrt{\vec{p}_{\vec{n}}^2 + m^2}. \quad (3.57)$$

The translational partition function is the sum over all the states, i.e., all triplets \vec{n} of positive integers,

$$Z_{\text{tr}} = \sum_{n_x=1}^{\infty} \sum_{n_y=1}^{\infty} \sum_{n_z=1}^{\infty} \exp[-\beta\epsilon(\vec{p}_{\vec{n}})]. \quad (3.58)$$

When we write this as

$$Z_{\text{tr}} = \frac{1}{(\Delta p)^3} \sum_{n_x=1}^{\infty} \Delta p \sum_{n_y=1}^{\infty} \Delta p \sum_{n_z=1}^{\infty} \Delta p \exp[-\beta\epsilon(\vec{p}_{\vec{n}})]. \quad (3.59)$$

When we take the limit $\Delta p \rightarrow 0$, this coincides with Riemann's definition of the integral

$$Z_{\text{tr}} = \frac{1}{(\Delta p)^3} \int_{\Delta p}^{\infty} dp_x \int_{\Delta p}^{\infty} dp_y \int_{\Delta p}^{\infty} dp_z \exp[-\beta\epsilon(\vec{p})]. \quad (3.60)$$

As $\Delta p \rightarrow 0$, the lower limits of the integrals approach zero, and therefore we have a three-dimensional momentum integral over momenta whose all components p_x , p_y and p_z are positive. Since the integrand does not depend on their signs, we can extend the integration over the whole momentum space by replacing

$$\int_0^{\infty} dp_x = \frac{1}{2} \int_{-\infty}^{\infty} dp_x. \quad (3.61)$$

Therefore, we find

$$Z_{\text{tr}} = \frac{1}{8(\Delta p)^3} \int d^3 p e^{-\beta \epsilon(\vec{p})} = \frac{L^3}{h^3} \int d^3 p e^{-\beta \epsilon(\vec{p})}. \quad (3.62)$$

Furthermore, since the integrand does not depend on the direction of \vec{p} , it is convenient to use spherical coordinates and integrate over the angles to obtain

$$Z_{\text{tr}} = \frac{4\pi L^3}{h^3} \int_0^\infty dp p^2 e^{-\beta \epsilon(p)}, \quad (3.63)$$

with $\epsilon(p) = \sqrt{p^2 + m^2}$. This expression is valid generally, but it is often useful to consider **non-relativistic** ($p \ll mc$) and **ultrarelativistic** ($p \gg mc$) limits in which it simplified.

In the non-relativistic case, the energy of the particle is $\epsilon(p) = p^2/2m$ (where we have dropped a constant term). We change the integration variable from p to ϵ by writing $p = \sqrt{2m\epsilon}$, and find $dp = \sqrt{m/2\epsilon} d\epsilon$. Using these, we obtain

$$Z_{\text{tr}} = \frac{4\pi L^3}{h^3} \int_0^\infty d\epsilon \sqrt{\frac{m}{2\epsilon}} 2m\epsilon e^{-\beta \epsilon} = \int_0^\infty d\epsilon \frac{2\pi V}{h^3} (2m)^{3/2} \epsilon^{1/2} e^{-\beta \epsilon}. \quad (3.64)$$

Comparing with Eq. (3.50), we can see that in the non-relativistic limit the density of translational states is

$$f_{\text{tr}}^{\text{NR}}(\epsilon) = \frac{2\pi V}{h^3} (2m)^{3/2} \epsilon^{1/2}. \quad (3.65)$$

In the ultrarelativistic case, the energy of the particle is $\epsilon(p) = cp$, and $dp = d\epsilon/c$. The partition function is

$$Z_{\text{tr}} = \frac{4\pi L^3}{h^3} \int_0^\infty \frac{d\epsilon \epsilon^2}{c^3} e^{-\beta \epsilon} = \int_0^\infty d\epsilon \frac{4\pi V}{c^3 h^3} \epsilon^2 e^{-\beta \epsilon}, \quad (3.66)$$

and we can read off the density of states of ultrarelativistic particles,

$$f_{\text{tr}}^{\text{UR}}(\epsilon) = \frac{4\pi V}{c^3 h^3} \epsilon^2. \quad (3.67)$$

As an example, let us consider the thermodynamics of a free non-relativistic particle. Let us assume that the particle has spin s , so that it has $g = 2s + 1$ internal spin states. The full single-particle partition function is then

$$Z_1 = g Z_{\text{tr}} = g \int_0^\infty \frac{2\pi V}{h^3} (2m)^{3/2} \epsilon^{1/2} e^{-\beta \epsilon}.$$

First, we calculate the integral (see Appendix A)

$$\int_0^\infty d\epsilon \epsilon^{1/2} e^{-\beta \epsilon} = \frac{1}{2} \sqrt{\frac{\pi}{\beta^3}}, \quad (3.68)$$

and obtain

$$Z_1 = \frac{2\pi V g}{h^3} (2m)^{3/2} \frac{1}{2} \sqrt{\frac{\pi}{\beta^3}} = gV \left(\frac{2\pi m}{\beta h^2} \right)^{3/2} = gV \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2}. \quad (3.69)$$

We can now immediately calculate the mean energy $\bar{\epsilon}$ of the particle using Eq. (3.27),

$$\bar{\epsilon} = -\frac{\partial \ln Z_1}{\partial \beta} = -\frac{\partial}{\partial \beta} \left[\ln gV - \frac{3}{2} \ln \frac{\beta h^2}{2\pi m} \right] = \frac{3}{2\beta} = \frac{3}{2} k_B T. \quad (3.70)$$

This is an example of a very general result in classical statistical mechanics, known as **classical equipartition of energy**: At temperature T , every degree of freedom has energy $k_B T/2$. In our case, the particle has three translational degrees of freedom, since it can move in three dimensions. Note that the equipartition does generally **not** apply to quantum mechanical systems.

3.5 Gas of Distinguishable Particles

We can now try to describe a gas that consists of a large number N of (non-relativistic) particles. Let us first assume that the particles are **distinguishable**, which would be true in classical mechanics. We will see that the results do not agree with basic observations, which shows that identical particles are fundamentally indistinguishable. **Therefore, this calculation does not describe nature!**

The partition function of N distinguishable particles is

$$Z_N = \sum_{r_1} \cdots \sum_{r_N} e^{-\beta \sum_{i=1}^N \epsilon_{r_i}}, \quad (3.71)$$

where r_i is the state of particle i . In the sum (3.71), the exponential factorises into terms each of which depends only on the state of one particle,

$$e^{-\beta \sum_{i=1}^N \epsilon_{r_i}} = \prod_{i=1}^N e^{-\beta \epsilon_{r_i}}. \quad (3.72)$$

Furthermore, when we consider the innermost sum in Eq. (3.71), only the factor $e^{-\beta \epsilon_{r_N}}$ depends on the summation index r_N . We can move all the other factors outside the summation. Similarly, we can move the first $N - 2$ factors outside the summation over r_{N-1} and so on, so that we have

$$Z_N = \sum_{r_1} e^{-\beta \epsilon_{r_1}} \sum_{r_2} e^{-\beta \epsilon_{r_2}} \cdots \sum_{r_N} e^{-\beta \epsilon_{r_N}}. \quad (3.73)$$

But this is just the product

$$Z_N = \prod_{i=1}^N \left(\sum_{r_i} e^{-\beta \epsilon_{r_i}} \right). \quad (3.74)$$

Finally, we note that all the factors in this product are identical, so that we have

$$Z_N = \left(\sum_r e^{-\beta \epsilon_r} \right)^N = Z_1^N. \quad (3.75)$$

From this, we would obtain the free energy

$$F_N = -k_B T \ln Z_N = -N k_B T \ln Z_1 = -N k_B T \left(\ln V + \frac{3}{2} \ln \frac{2\pi m k_B T}{h^2} \right). \quad (3.76)$$

From thermodynamics, we know that the free energy should be an extensive quantity, so that if we double the system size, the free energy doubles, $F_N \rightarrow 2F_N$. However, the expression in Eq. (3.76) is not extensive, since if we double the system size ($N \rightarrow 2N$, $V \rightarrow 2V$), we find

$$F_N \rightarrow -2N k_B T \left(\ln 2V + \frac{3}{2} \ln \frac{2\pi m k_B T}{h^2} \right) = 2F_N - 2N k_B T \ln 2 \neq 2F_N. \quad (3.77)$$

This shows that we must have made done something wrong, and indeed we have. We assumed implicitly that the particles are distinguishable so that we can label them by i , but that is not true.

4 Variable Particle Number

4.1 Grand Canonical Ensemble

Distinguishable and indistinguishable particles were already discussed in Section 2.2. As stated there, it is often most convenient to represent the microstate in terms of occupation numbers n_r . The total number of particles is

$$N = \sum_r n_r, \quad (4.1)$$

where the sum goes over all single-particle states r , and the total energy is

$$U = \sum_r \epsilon_r n_r, \quad (4.2)$$

where ϵ_r is the energy of the single-particle state r . In the partition function, we have to sum over all different microstates, and for N indistinguishable particles this means

$$Z_N = \left(\sum_{n_0} \sum_{n_1} \cdots \right) e^{-\beta \sum_{r=0}^{\infty} n_r \epsilon_r}, \quad \text{with } \sum_r n_r = N, \quad (4.3)$$

rather than Eq. (3.71).

It is quite clear that it is difficult to calculate the sum in Eq. (4.3) in practice because of the constraint $\sum_r n_r = N$. The situation is very similar to the microcanonical ensemble, in which calculations are made difficult by the energy constraint

$$\sum_i \epsilon_{r_i} = U. \quad (4.4)$$

Then, the problem was solved by switching to the canonical ensemble in which U is allowed to vary. In the same way, we can make N variable.

Recall the derivation of the canonical ensemble in Section 3.2. There we allowed system 1 to exchange energy with the much bigger system 2, which acted as a heat bath. Now, let us allow them to exchange particles, too. The macrostates that we consider are parameterised by the energies U_1, U_2 and particle numbers N_1, N_2 of the two systems. The total energy $U = U_1 + U_2$ and the total particle number $N = N_1 + N_2$ are fixed, but the energies and particle numbers of the individual systems are allowed to vary.

The entropy S of the whole system is again a sum of the entropies S_1 and S_2 of the two subsystems,

$$S(U_1, N_1; U_2, N_2) = S_1(U_1, N_1) + S_2(U_2, N_2). \quad (4.5)$$

According to the fundamental hypothesis, the probability of the macrostate $(U_1, N_1; U_2, N_2)$ is proportional to its multiplicity, and using Boltzmann's definition of entropy (2.18), we can write

$$\begin{aligned} p(U_1, N_1; U_2, N_2) &\propto e^{S(U_1, N_1; U_2, N_2)/k_B} = e^{S_1(U_1, N_1)/k_B} e^{S_2(U_2, N_2)/k_B} \\ &= \Omega_1(U_1, N_1) e^{S_2(U_2, N_2)/k_B}, \end{aligned} \quad (4.6)$$

where $\Omega_1(U_1, N_1)$ is the number of microstates of system 1 that have energy U_1 and particle number N_1 .

It again turns out to be useful to Taylor expand

$$S_2(U, N) = S_2(U_2 + U_1, N_2 + N_1) \approx S_2(U_2, N_2) + \frac{\partial S_2}{\partial U_2} U_1 + \frac{\partial S_2}{\partial N_2} N_1 + \dots \quad (4.7)$$

In analogy with the definition of temperature (2.19) we can define a new quantity called the **chemical potential** and denoted by μ as

$$\frac{\mu}{T} = -\frac{\partial S_2}{\partial N_2}. \quad (4.8)$$

As with the temperature, we assume that μ is a constant, because system 2 is much larger than system 1. The probability that system 1 has energy U_1 and particle number N_1 is therefore

$$p(U_1, N_1) = \frac{\Omega_1(U_1, N_1)}{\mathcal{Z}} \exp\left(-\frac{1}{k_B T} U_1 + \frac{\mu}{k_B T} N_1\right), \quad (4.9)$$

where \mathcal{Z} is a normalisation constant known as the **grand partition function**. We can obtain the probability of a single microstate by dividing this by $\Omega_1(U_1, N_1)$.

Because all references to system 2 have disappeared, we also drop the subscript 1, and simply say that the result is valid for any system that can exchange energy and particles with its environment. Microstates with any energy and particle number are allowed, and the probability that the system is in a microstate α with energy E_α and particle number N_α is

$$p_\alpha = \frac{1}{\mathcal{Z}} e^{-(E_\alpha - \mu N_\alpha)/k_B T}. \quad (4.10)$$

This defines the **grand canonical ensemble**.

From the condition that the sum of the probabilities of all microstates have to be 1,

$$1 = \sum_{\alpha} p_\alpha = \frac{1}{\mathcal{Z}} \sum_{\alpha} e^{-(E_\alpha - \mu N_\alpha)/k_B T}, \quad (4.11)$$

we obtain an explicit expression for the grand partition function,

$$\mathcal{Z} = \sum_{\alpha} e^{-(E_\alpha - \mu N_\alpha)/k_B T}. \quad (4.12)$$

The connection with thermodynamics is provided by the **grand potential** Φ , defined as⁴

$$\Phi = -k_B T \ln \mathcal{Z}. \quad (4.13)$$

It can be related to other thermodynamic variables as (Problem Sheet 3)

$$\Phi = U - TS - \mu N. \quad (4.14)$$

From this it follows that the entropy S , particle number N and pressure P of the system are given by the derivatives

$$S = -\left(\frac{\partial \Phi}{\partial T}\right)_{V, \mu}, \quad N = -\left(\frac{\partial \Phi}{\partial \mu}\right)_{T, V}, \quad P = -\left(\frac{\partial \Phi}{\partial V}\right)_{T, \mu}. \quad (4.15)$$

⁴Different authors use different symbols for the grand potential. A common choice is Ω , but it risks confusion with the multiplicity. Another popular one is Φ_G , but to simplify the notation, I have dropped the subscript.

ensemble	fixed variables	probability of a microstate	link with thermodynamics
Microcanonical	U, N, V	$p_\alpha = 1/\Omega$	$S = k_B \ln \Omega$
Canonical	N, V	$p_\alpha = (1/Z)e^{-\beta E_\alpha}$	$F = -k_B T \ln Z$
Grand canonical	V	$p_\alpha = (1/\mathcal{Z})e^{-\beta(E_\alpha - \mu N_\alpha)}$	$\Phi = -k_B T \ln \mathcal{Z}$

Table 3: Summary of the different ensembles.

It is important to realise that in this description, the particle number N is not fixed but it fluctuates, just like energy U in the canonical ensemble. The chemical potential μ determines the mean particle number \bar{N} , just like the temperature determines the mean energy \bar{E} . In essence, the chemical potential measures the willingness of the environment to give more particles to system 1. If μ is high, N will also be high. One can also show that

$$\frac{\Delta N}{N} \propto \frac{1}{\sqrt{N}}, \quad (4.16)$$

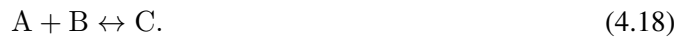
and therefore the grand canonical ensemble actually agrees with the other ensembles in the thermodynamic limit $N \rightarrow \infty$. One is therefore free to choose whichever ensemble is the most convenient for the given problem.

Chemical potential μ may seem like an abstract quantity that is hard to visualise, so it is useful to keep in mind that its relation to particle number N is the same as the relation of temperature T to energy U . When two systems are put in thermal contact, heat flows from higher to lower temperature. When two systems are allowed to exchange particles, particles flow from higher to lower chemical potential. In equilibrium, the chemical potentials of the two systems are equal.

Chemical potential can also be used to describe reactions in which particle species change, such as chemical reactions. Let us, for instance, consider three particle species A, B and C and denote the particle numbers of species X by N_X . By definition, the chemical potential of species X is

$$\mu_X = -T \left(\frac{\partial S}{\partial N_X} \right)_{V, U}. \quad (4.17)$$

Let us now assume that A and B can react and produce C, in the reaction



This reaction reduces N_A and N_B by one and increases N_C by one. If we denote the initial particle number of species X by N_X^0 , then after Δ reactions, the particle numbers are

$$\begin{aligned} N_A &= N_A^0 - \Delta, \\ N_B &= N_B^0 - \Delta, \\ N_C &= N_C^0 + \Delta. \end{aligned} \quad (4.19)$$

The reactions continue until the system reaches equilibrium. To find this state, we need to find the value of Δ that maximises the entropy of the system, i.e.,

$$\frac{\partial S}{\partial \Delta} = 0. \quad (4.20)$$

Using Eq. (4.19), we find

$$0 = \frac{\partial S}{\partial \Delta} = -\frac{\partial S}{\partial N_A} - \frac{\partial S}{\partial N_B} + \frac{\partial S}{\partial N_C} = \frac{\mu_A + \mu_B - \mu_C}{T}. \quad (4.21)$$

Therefore, the equilibrium condition is

$$\mu_A + \mu_B = \mu_C. \quad (4.22)$$

More generally, for an arbitrary reaction to be in equilibrium, the total chemical potentials on the left and right hand sides of the equation have to match.

4.2 Bose-Einstein and Fermi-Dirac Distributions

The grand partition function of a quantum gas is given by the sum over the occupation number n_r of all single-particle states r ,

$$\mathcal{Z} = \left(\sum_{n_0} \sum_{n_1} \dots \right) e^{-\beta \sum_r n_r (\epsilon_r - \mu)}. \quad (4.23)$$

Just like Eq. (3.71), this factorises,

$$\mathcal{Z} = \left(\sum_{n_0} e^{-\beta n_0 (\epsilon_0 - \mu)} \right) \left(\sum_{n_1} e^{-\beta n_1 (\epsilon_1 - \mu)} \right) \dots = \prod_r \left(\sum_{n_r} e^{-\beta n_r (\epsilon_r - \mu)} \right) \equiv \prod_r \mathcal{Z}_r, \quad (4.24)$$

where

$$\mathcal{Z}_r = \sum_{n_r} e^{-\beta n_r (\epsilon_r - \mu)} \quad (4.25)$$

can be interpreted as the partition function of the single-particle state r .

From Eq. (4.25) we can see that the probability that the single-particle state r has occupation number n_r is

$$p_{r,n_r} = \frac{e^{-\beta (\epsilon_r - \mu) n_r}}{\mathcal{Z}_r}. \quad (4.26)$$

As a check, we can see that according to this, the probability that the system is in a given microstate $\alpha = [n_0, n_1, \dots]$ is the joint probability of state 0 having occupation number n_0 , state 1 having occupation number n_1 etc., i.e.,

$$p_\alpha = \prod_r p_{r,n_r} = \prod_r \frac{e^{-\beta (\epsilon_r - \mu) n_r}}{\mathcal{Z}_r} = \frac{e^{-\beta \sum_r (\epsilon_r - \mu) n_r}}{\prod_r \mathcal{Z}_r} = \frac{e^{-\beta (E_\alpha - \mu N_\alpha)}}{\mathcal{Z}}, \quad (4.27)$$

in agreement with Eq. (4.10).

For most thermodynamic quantities, it is enough to know the mean occupation numbers \bar{n}_r of states r . They are collectively known as the **distribution** of particles. The mean occupation number \bar{n}_r of state r is given by

$$\bar{n}_r = \sum_{n_r} n_r p_{r,n_r} = \frac{1}{\mathcal{Z}_r} \sum_{n_r} n_r e^{-\beta (\epsilon_r - \mu) n_r}. \quad (4.28)$$

We note that

$$\frac{\partial \mathcal{Z}_r}{\partial \epsilon_r} = \sum_{n_r} \frac{\partial e^{-\beta (\epsilon_r - \mu) n_r}}{\partial \epsilon_r} = \sum_{n_r} (-\beta n_r) e^{-\beta (\epsilon_r - \mu) n_r} = -\beta \sum_{n_r} n_r e^{-\beta (\epsilon_r - \mu) n_r}, \quad (4.29)$$

so that we can write

$$\bar{n}_r = -\frac{1}{\beta} \frac{1}{\mathcal{Z}_r} \frac{\partial \mathcal{Z}_r}{\partial \epsilon_r} = -\frac{1}{\beta} \frac{\partial \ln \mathcal{Z}_r}{\partial \epsilon_r}. \quad (4.30)$$

The range of the summation in Eq. (4.25) depends on the nature of the particles. For **bosons**, we have

$$\mathcal{Z}_r = \sum_{n_r=0}^{\infty} e^{-\beta n_r(\epsilon_r - \mu)} = \sum_{n_r=0}^{\infty} \left(e^{-\beta(\epsilon_r - \mu)} \right)^{n_r}. \quad (4.31)$$

This is a geometric series and can be summed exactly (assuming $\mu < \epsilon_r$),

$$\mathcal{Z}_r = \frac{1}{1 - e^{-\beta(\epsilon_r - \mu)}}. \quad (4.32)$$

We can then compute the mean occupation number using Eq. (4.30),

$$\begin{aligned} \bar{n}_r &= -\frac{1}{\beta} \frac{\partial}{\partial \epsilon_r} \ln \frac{1}{1 - e^{-\beta(\epsilon_r - \mu)}} = \frac{1}{\beta} \frac{\partial}{\partial \epsilon_r} \ln \left(1 - e^{-\beta(\epsilon_r - \mu)} \right) \\ &= \frac{1 - (-\beta)e^{-\beta(\epsilon_r - \mu)}}{\beta (1 - e^{-\beta(\epsilon_r - \mu)})} = \frac{1}{e^{\beta(\epsilon_r - \mu)} - 1}. \end{aligned} \quad (4.33)$$

This is known as the **Bose-Einstein distribution**, which is often abbreviated as BE. Because it only depends on the energy ϵ of the state, we can drop the index r and write⁵

$$\bar{n}_{\text{BE}}(\epsilon) = \frac{1}{e^{\beta(\epsilon - \mu)} - 1}. \quad (4.34)$$

For **fermions**, the calculation of \mathcal{Z}_r is even simpler,

$$\mathcal{Z}_r = \sum_{n_r=0}^1 e^{-\beta n_r(\epsilon_r - \mu)} = 1 + e^{-\beta(\epsilon_r - \mu)}. \quad (4.35)$$

The mean occupation number is then

$$\begin{aligned} \bar{n}_r &= -\frac{1}{\beta} \frac{\partial}{\partial \epsilon_r} \ln \left(1 + e^{-\beta(\epsilon_r - \mu)} \right) = -\frac{1}{\beta} \frac{(-\beta)e^{-\beta(\epsilon_r - \mu)}}{1 + e^{-\beta(\epsilon_r - \mu)}} \\ &= \frac{1}{e^{\beta(\epsilon_r - \mu)} + 1}. \end{aligned} \quad (4.36)$$

This is known as the **Fermi-Dirac distribution** (abbreviated FD). Again, we write it as a function of energy,

$$\bar{n}_{\text{FD}}(\epsilon) = \frac{1}{e^{\beta(\epsilon - \mu)} + 1}. \quad (4.37)$$

We can summarize Eqs. (4.34) and (4.37) in one expression,

$$\bar{n}(\epsilon) = \frac{1}{e^{\beta(\epsilon - \mu)} \pm 1}, \quad (4.38)$$

where $+$ corresponds to the Fermi-Dirac and $-$ to the Bose-Einstein distribution. By looking at the asymptotic behaviour, one can draw the following conclusions:

⁵Sturge and Trevena use the symbol f for Bose-Einstein and Fermi-Dirac distributions.

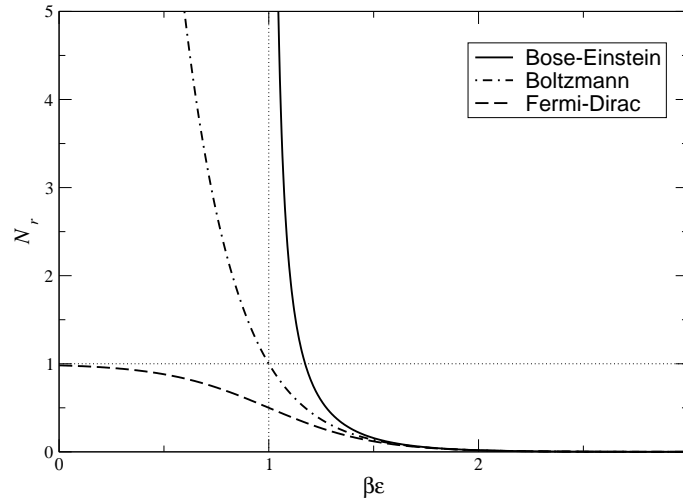


Figure 3: The shapes of Bose-Einstein, Fermi-Dirac and the classical Boltzmann distribution for $\beta = 4$ and $\beta\mu = 1$.

- When $\beta(\epsilon - \mu) \gg 1$ (which is valid for any ϵ if $\beta\mu \ll -1$), the term ± 1 becomes negligible, and both distributions approach

$$\bar{n}(\epsilon) \rightarrow e^{\beta\mu} e^{-\beta\epsilon} \propto e^{-\beta\epsilon}, \quad (4.39)$$

which agrees with the Boltzmann distribution obeyed by distinguishable particles. In that case, $\bar{n}_r = N p_r = (N/Z) e^{-\beta(\epsilon_r)}$. In this limit, the quantum gas approaches a classical ideal gas, but without the extensivity problem (3.77) (see Classwork 2). In this same limit, the mean occupation number is also very low, $\bar{n}_r \ll 1$, and therefore it is natural that the difference between bosons and fermions disappears.

- In the Fermi-Dirac distribution, the mean occupation number $\bar{n}_{\text{FD}}(\epsilon)$ approaches one when $\epsilon \ll \mu$. Because one is the maximum number of particles in one state, this means that states with low energies are **fully occupied**.
- In the Bose-Einstein distribution, the mean occupation number $\bar{n}_{\text{BE}}(\epsilon)$ diverges when $\epsilon \rightarrow \mu$. This means that for all occupation numbers to be finite, μ must be less than the ground state energy,

$$\mu < \epsilon_0. \quad (4.40)$$

The shapes of the distributions are shown in Fig. 3.

5 Bosonic Gases

5.1 Black-body Radiation

Black-body radiation is thermal electromagnetic radiation, which we can think of as photons in thermal equilibrium. Photons are bosons (since they have spin 1), so they are described by the Bose-Einstein distribution. Because photons are massless, they are always ultrarelativistic. For a given momentum, there are two independent polarisation states (for instance vertical and horizontal).

The chemical potential μ of a photon gas vanishes, because the number of photons is not conserved. To see this concretely, consider a process in which a hydrogen atom absorbs a photon of energy 12.1 eV. This takes the atom from its ground state $n = 1$ to an excited state $n = 3$. It then decays first to state $n = 2$ and then back to ground state $n = 1$, emitting two photons of energies 1.9 eV and 10.2 eV. If we represent the photon by the symbol γ and the hydrogen atom by H, we can write this reaction as



For this reaction to be in equilibrium, we must have

$$\mu_{\text{H}} + \mu_{\gamma} = \mu_{\text{H}} + 2\mu_{\gamma}, \quad (5.2)$$

which implies $\mu_{\gamma} = 0$.

Another way of understanding why $\mu = 0$ is that the chemical potential tells how much bias towards higher or lower particle number there is due to the interaction with the environment. As the particle number is not conserved, it will always reach the same equilibrium value even if we try to pump more photons into the system.

In fact, since the photon number cannot be fixed, there is no constraint in canonical partition function Eq. (4.3). Without the constraint, the sum in Eq. (4.3) is the same as that in Eq. (4.12) with $\mu = 0$. One can therefore equally well say that we are using the canonical ensemble Eq. (4.3) but ignoring the particle number, or that we use the grand canonical ensemble with $\mu = 0$. We follow the latter interpretation.

Because $\mu = 0$, the mean occupation number \bar{n}_r of single-particle state r is

$$\bar{n}_r = \frac{1}{e^{\beta\epsilon_r} - 1}. \quad (5.3)$$

As usual, the total internal energy is

$$U = \sum_r \bar{n}_r \epsilon_r = \sum_r \frac{\epsilon_r}{e^{\beta\epsilon_r} - 1}. \quad (5.4)$$

As in Section 3.4, we replace the sum over r by an integral over energy using the ultrarelativistic density of states (3.67). Photons have two internal states, corresponding to two polarizations, so there is a degeneracy factor $g = 2$, and we obtain

$$U = \int_0^{\infty} d\epsilon f(\epsilon) \frac{\epsilon}{e^{\beta\epsilon} - 1} = g \int_0^{\infty} d\epsilon f_{\text{tr}}^{\text{UR}}(\epsilon) \frac{\epsilon}{e^{\beta\epsilon} - 1} = \frac{8\pi V}{h^3 c^3} \int_0^{\infty} d\epsilon \epsilon^2 \frac{\epsilon}{e^{\beta\epsilon} - 1}. \quad (5.5)$$

A more useful quantity is the energy density $u = U/V$, and using $\epsilon = \hbar\omega = h\omega/2\pi$, we can write it as an integral over frequency ω ,

$$u = \frac{8\pi}{h^3 c^3} \int_0^{\infty} \frac{d\epsilon \epsilon^3}{e^{\beta\epsilon} - 1} = \int_0^{\infty} d\omega \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\beta\hbar\omega} - 1} \equiv \int_0^{\infty} d\omega u_s(\omega), \quad (5.6)$$

where the **spectral energy density** $u_s(\omega)$ gives the contribution to the energy density from a given frequency ω , and is

$$u_s(\omega) = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\beta\hbar\omega} - 1}. \quad (5.7)$$

This is the **Planck radiation law**, and it gives the spectrum of thermal radiation.

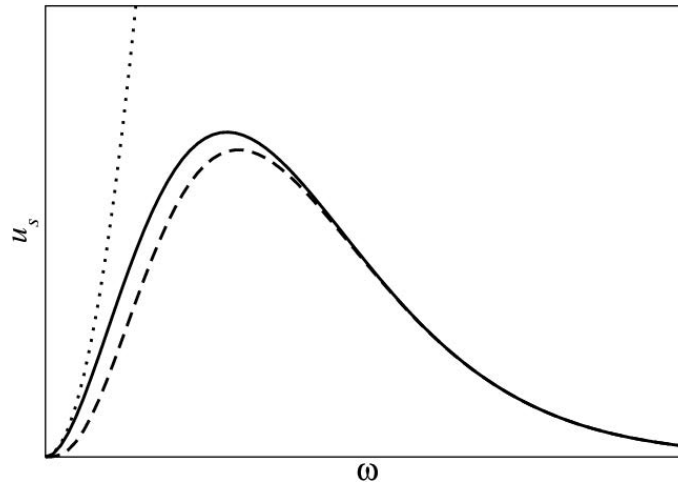


Figure 4: The Planck radiation law (5.7) [solid curve] agrees with the Wien law (5.8) [dashed] at high frequencies and with the Rayleigh law (5.9) [dotted] at low frequencies.

Historically, Planck introduced Eq. (5.7) as an interpolating formula between the empirical Wien law

$$u_s(\omega) = \frac{\hbar\omega^3}{\pi^2c^3}e^{-\beta\hbar\omega}, \quad (5.8)$$

which agreed with experiments, and the theoretical prediction

$$u_s(\omega) = \frac{\omega^2}{\pi^2c^3}k_B T, \quad (5.9)$$

which Rayleigh had derived from classical electrodynamics (see Fig. 4). Interestingly, the Rayleigh result showed even without any experiments that classical physics cannot describe thermal radiation, since Eq. (5.9) would give a divergent result for the energy density u .

We can do the integral in Eq. (5.6) using the standard integral (Appendix A)

$$\int_0^\infty \frac{dx x^3}{e^x - 1} = \frac{\pi^4}{15}, \quad (5.10)$$

and we find

$$u = \frac{\pi^2 k_B^4}{15 \hbar^3 c^3} T^4 \equiv \frac{4\sigma}{c} T^4, \quad (5.11)$$

where

$$\sigma = \frac{\pi^2 k_B^4}{60 \hbar^3 c^2} \approx 5.7 \times 10^{-8} \text{W/m}^2 \text{K}^4 \quad (5.12)$$

is the **Stefan-Boltzmann constant**. The result (5.11) is known as the **Stefan-Boltzmann law**, and gives the energy density of thermal radiation as a function of temperature.

In practice, the radiation from many hot sources is not fully thermal and is therefore not in full agreement with the Planck law. However, the cosmic microwave background radiation is a notable exception. It is radiation emitted when the universe became transparent as it was approximately 300000 years old. The spectrum of this radiation was measured by the COBE satellite in 1992 (see Fig. 5), and it agrees amazingly well with the Planck law, much better than any other radiation. This was a definite proof for the hot Big Bang theory in cosmology, and is also a beautiful demonstration of the validity of both quantum mechanics and statistical physics.

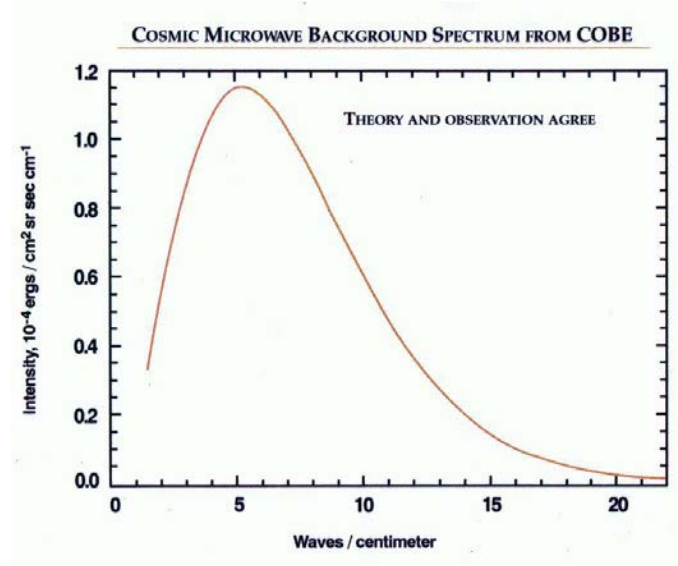


Figure 5: The spectrum of the cosmic microwave background as measured by the COBE satellite <http://lambda.gsfc.nasa.gov/product/cobe/>. The plot shows both the theoretical prediction and the data, but the agreement is so good that it is not possible to see the difference.

5.2 Bose-Einstein Condensation

As another example of a bosonic gas, let us consider N non-relativistic bosonic atoms at temperature T . According to Eq. (4.33), the mean occupation number of state r is

$$\bar{n}_r = \frac{1}{e^{\beta(\epsilon_r - \mu)} - 1}. \quad (5.13)$$

The total number of particles is therefore

$$N = \sum_r \bar{n}_r. \quad (5.14)$$

Since N is known, we can use this to fix the value of the chemical potential μ , but we will first have to be able to calculate the sum.

To do this, we follow Section. 3.4, and replace the sum by an integral

$$N \rightarrow \int_0^\infty d\epsilon f(\epsilon) \bar{n}_{\text{BE}}(\epsilon). \quad (5.15)$$

For spin-0 atoms, the degeneracy factor is zero, $g = 0$, and the the density of states is given by Eq. (3.65). Consequently, we find

$$N = \frac{2\pi V}{h^3} (2m)^{3/2} \int_0^\infty d\epsilon \epsilon^{1/2} \bar{n}_{\text{BE}}(\epsilon). \quad (5.16)$$

The chemical potential μ of a bosonic gas is limited by Eq. (4.40) to be negative, $\mu < 0$, assuming that the ground state energy is zero. This leads to an upper bound for the occupation numbers,

$$\bar{n}(\epsilon) < \frac{1}{e^{\beta\epsilon} - 1}, \quad (5.17)$$

and consequently for the total number of particles,

$$N < \frac{2\pi V}{h^3} (2m)^{3/2} \int_0^\infty \frac{d\epsilon \epsilon^{1/2}}{e^{\beta\epsilon} - 1}. \quad (5.18)$$

We can do this integral by changing the integration variable to $x = \beta\epsilon$,

$$\int_0^\infty \frac{d\epsilon \epsilon^{1/2}}{e^{\beta\epsilon} - 1} = \beta^{-3/2} \int_0^\infty \frac{dx x^{1/2}}{e^x - 1} = \beta^{-3/2} \frac{\sqrt{\pi}}{2} \zeta\left(\frac{3}{2}\right), \quad (5.19)$$

where we have used a standard integral from Appendix A, and the numerical value of the zeta function is $\zeta(3/2) \approx 2.61$. Therefore, we have the upper limit

$$N < V \left(\frac{2\pi m}{\beta h^2}\right)^{3/2} \zeta\left(\frac{3}{2}\right) = V \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} \zeta\left(\frac{3}{2}\right) \equiv N_{\max}(T), \quad (5.20)$$

where we have chosen to denote the upper limit by $N_{\max}(T)$.

Interestingly, if we keep N and V constant and decrease T , i.e., cool the system down in a rigid container, we will eventually reach a point where $N = N_{\max}(T)$, and then Eq. (5.20) seems to say that we cannot cool the system down any further. The system seems to have a minimum temperature T_B , which is determined by the equation $N = N_{\max}(T_B)$ and has the value

$$T_B = \frac{h^2}{2\pi m k_B} \left(\frac{n}{\zeta(3/2)}\right)^{2/3}, \quad (5.21)$$

where $n = N/V$ is the number density. This does not seem to make sense, since surely we should be able to cool the system down to arbitrarily low temperatures, at least in principle.

And indeed, we had made a mistake by assuming that we can replace the sum by an integral (5.15). Since $\lim_{\epsilon \rightarrow 0} f(\epsilon) = 0$, the integral gives no weight to the ground state $r = 0$ of the system, since it has $\epsilon = 0$. If there happen to be a large number of particles in the ground state, this approximation fails. To find out if this is the case, we count the ground state separately, and replace Eq. (5.15) by

$$N \rightarrow \bar{n}_0 + \int_0^\infty d\epsilon f(\epsilon) \bar{n}(\epsilon). \quad (5.22)$$

When $T > T_B$, we already had a consistent picture, so we do not expect this to change anything. However, at $T < T_B$, we assume that all the “extra” particles are in the ground state. That is, we write

$$\begin{aligned} \bar{n}_0 &= N - \frac{2\pi V}{h^3} (2m)^{3/2} \int_0^\infty \frac{d\epsilon \epsilon^{1/2}}{e^{\beta(\epsilon-\mu)} - 1} \\ &\approx N - N_{\max} = N - V \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} \zeta\left(\frac{3}{2}\right) = N \left[1 - \left(\frac{T}{T_B}\right)^{3/2}\right], \end{aligned} \quad (5.23)$$

where on the second line, we have assumed that $\mu = 0$ in the integral in Eq. (5.22)

Directly from the Bose-Einstein distribution (4.33) we also have another expression for \bar{n}_0 ,

$$\bar{n}_0 = \frac{1}{e^{-\beta\mu} - 1}, \quad (5.24)$$

which we can solve for μ and find

$$\mu = -\frac{1}{\beta} \ln\left(1 + \frac{1}{\bar{n}_0}\right). \quad (5.25)$$

Since $\bar{n}_0 \propto N$ is generally large, we can Taylor expand this,

$$\mu \approx -\frac{1}{\beta\bar{n}_0} = -\left[1 - \left(\frac{T}{T_B}\right)^{3/2}\right]^{-1} \frac{1}{\beta N}, \quad (5.26)$$

where we have also used Eq. (5.23).

Now, one has to ask how well the integral approximates the other low-lying energy states, given that it failed so miserably with the ground states. Would the first excited state $r = 1$ also have a large occupation number \bar{n}_1 ?

To answer this, let us first calculate its energy. According to Eq. (3.56), the momenta of the particles are quantised as

$$\vec{p} = \frac{h}{2L}\vec{n}, \quad (5.27)$$

where \vec{n} is a vector of three positive integers. The ground state corresponds to $\vec{n}_0 = (1, 1, 1)$. The lowest three excited states all have the same energy, and correspond to different orientations of $\vec{n}_0 = (2, 1, 1)$. As we have chosen to normalise the ground state energy to zero, the energy of the first excited state is

$$\epsilon_1 = \frac{p_1^2 - p_0^2}{2m} = \frac{h^2}{8mL^2} (\vec{n}_1^2 - \vec{n}_0^2) = \frac{h^2}{8mL^2} (6 - 3) = \frac{3h^2}{8mL^2}. \quad (5.28)$$

Thus $\epsilon_1 \propto L^{-2} = V^{-2/3}$, and in the limit of very large volume, this is always much larger than $|\mu| \propto N^{-1} \propto V^{-1}$. Thus,

$$\epsilon_1 \gg |\mu|. \quad (5.29)$$

Therefore, we can well approximate the mean occupation number of the first excited state as

$$\bar{n}_1 = \frac{1}{e^{\beta(\epsilon_1 - \mu)} - 1} \approx \frac{1}{e^{\beta\epsilon_1} - 1}. \quad (5.30)$$

From this we can immediately conclude that it was consistent to take $\mu = 0$ in Eq. (5.23).

Further, if V is large, $\beta\epsilon_1 \propto V^{-2/3}$ is always very small, and we can Taylor expand the exponential in Eq. (5.30),

$$\bar{n}_1 \approx \frac{1}{1 + \beta\epsilon_1 - 1} = \frac{1}{\beta\epsilon_1} \ll \frac{1}{\beta|\mu|} = \bar{n}_0, \quad (5.31)$$

where we have used Eq. (5.29). This means that the mean occupation number of the first excited state is much less than that of the ground state. Since any higher states have even smaller mean occupation numbers, we conclude that it is enough to treat the ground state separately, and all excited states can be well approximated by the integral.

Thus, we have seen that at low temperatures $T < T_B$, a macroscopic number of the particles are in the ground state. This phenomenon is known as **Bose-Einstein condensation** and was predicted by Einstein in 1924. A Bose-Einstein condensate has curious properties, since the particles in the ground state carry no energy and do not move (since they have a stationary wave function). As a consequence, they do not contribute to the internal energy density, pressure, viscosity or many other macroscopic characteristics of the system. This means that at constant temperature and volume, these properties do not change if one adds more particles into the system. They are therefore independent of the particle density. For example, the pressure of a Bose-Einstein condensate does not increase when it is compressed.

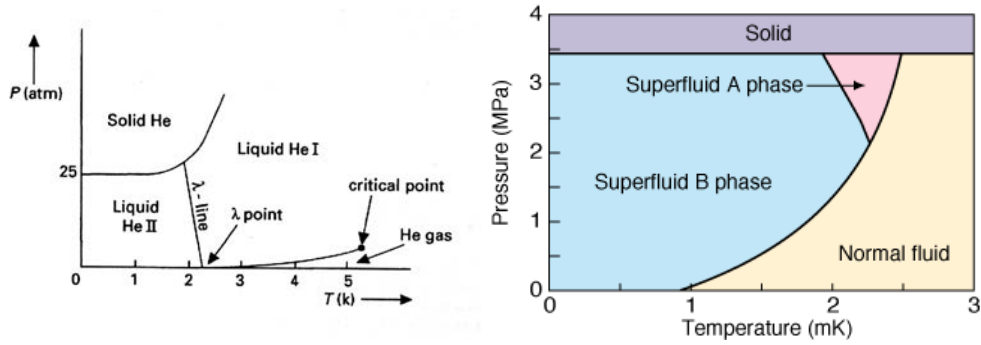


Figure 6: The phase diagrams of helium isotopes ${}^4\text{He}$ (left) and ${}^3\text{He}$ (right). Images: Trevena, <http://l11.tkk.fi/research/theory/he3.html>

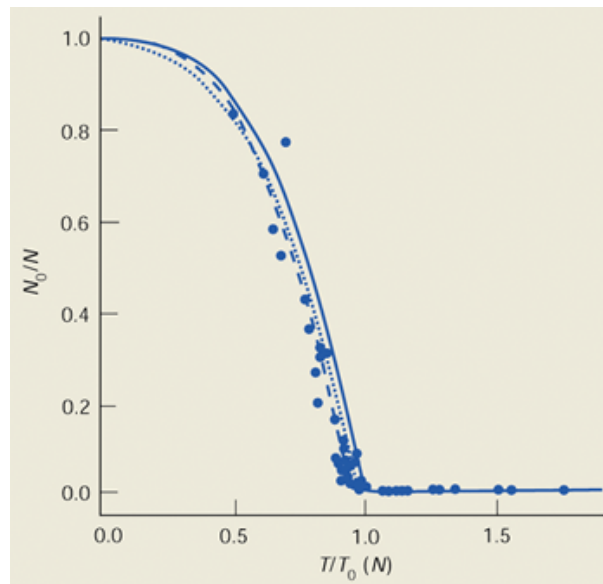


Figure 7: The measured occupation number of the ground state in ${}^{23}\text{Na}$. The solid curve is the prediction (5.23). Image: <http://physicsweb.org/articles/world/10/3/3>

For decades, the only example of a system that could be thought of as a Bose-Einstein condensate was liquid ^4He . The atoms of this helium isotope contain an even number of elementary particles, and are therefore bosons. At a temperature of around 2K, it has a phase transition and becomes superfluid (see Fig. 6). This is near the predicted Bose-Einstein condensation temperature, and the superfluid phase has some similar characteristics to Bose-Einstein condensates. However, since it is a liquid, it is clear the one cannot really assume the atoms are weakly interacting as in a Bose-Einstein condensate. Interestingly, though, the ^3He isotope, which is a fermion, also becomes superfluid but only at a much lower temperature $\approx 2\text{mK}$. On one hand the huge difference in the critical temperatures shows that the bosonic nature of ^4He plays an important role in superfluidity, but on the other, it does not seem absolutely necessary since even fermionic ^3He exhibits superfluidity.

Bose-Einstein condensation was finally confirmed unambiguously in 1995 by groups at Colorado and MIT. They used very dilute atomic gases (^{23}Na and ^{87}Rb), so that the interactions between the particles were extremely weak. The gas was trapped in a magnetic trap and cooled down to μK temperatures. When the trap was removed, the gas expanded, and the number of particles in the ground state could be measured from the way it expanded. As Fig. 7 shows the result agrees very well with the theoretical prediction (5.23).

6 Fermionic Gases

6.1 Degenerate Fermion Gas

The behaviour of fermion gases at low temperature is very different from bosons. They are described by the Fermi-Dirac distribution (4.36),

$$\bar{n}_{\text{FD}}(\epsilon) = \frac{1}{e^{\beta(\epsilon-\mu)} + 1}. \quad (6.1)$$

At low temperatures, $\beta \rightarrow \infty$, and therefore the exponent goes to either plus or minus infinity, depending on the signs of $(\epsilon - \mu)$, therefore we find

$$\bar{n}_{\text{FD}} \xrightarrow{\beta \rightarrow \infty} \theta(\mu - \epsilon) \equiv \begin{cases} 1 & \text{if } \epsilon < \mu, \\ 0 & \text{if } \epsilon > \mu, \end{cases} \quad (6.2)$$

where we have used the **step function** $\theta(x)$ which is defined as

$$\theta(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x < 0. \end{cases} \quad (6.3)$$

This means that all states with $\epsilon < \mu$ are filled, and there are no particles in any of the higher states. In this case, the gas is said to be **degenerate**.⁶ The zero temperature value of μ is called the **Fermi Energy** and denoted by ϵ_F ,

$$\epsilon_F = \mu|_{T=0}. \quad (6.4)$$

Sometimes one also talks about the **Fermi temperature**

$$T_F = \frac{\epsilon_F}{k_B}, \quad (6.5)$$

which has the significance that at $T < T_F$ the description as a degenerate fermion gas starts to be a good approximation. Unlike in the case of bosons, there is no sharp phase transition.

⁶Do not confuse with the other meaning of the term degenerate discussed in Section 3.4.

The total particle number is

$$N = \sum_r \bar{n}_r \rightarrow \int_0^\infty d\epsilon f(\epsilon) \bar{n}_{\text{FD}}(\epsilon). \quad (6.6)$$

6.2 Non-Relativistic Degenerate Electron Gas

Let us now consider a degenerate electron gas. Electrons are fermions with spin $1/2$, so the degeneracy is $g = 2$. If the Fermi energy is much less than the rest energy of the particle, $\epsilon_F \ll mc^2$, the electrons are non-relativistic, and the density of states is given by (3.65),

$$f(\epsilon) = 2f_{\text{tr}}^{\text{NR}}(\epsilon) = \frac{4\pi V}{h^3} (2m)^{3/2} \epsilon^{1/2}. \quad (6.7)$$

Using Eq. (6.2) we find that in the low-temperature limit the particle number becomes

$$N \xrightarrow{T \rightarrow 0} \int_0^{\epsilon_F} d\epsilon f(\epsilon) = \frac{4\pi V}{h^3} (2m)^{3/2} \int_0^{\epsilon_F} d\epsilon \epsilon^{1/2} = \frac{4\pi V}{h^3} (2m)^{3/2} \frac{2}{3} \epsilon_F^{3/2} = \frac{8\pi V}{3} \left(\frac{2m\epsilon_F}{h^2} \right)^{3/2}. \quad (6.8)$$

In practice, we would usually know N (or $n = N/V$), so it makes more sense to invert this result, and write ϵ_F as a function of n ,

$$\epsilon_F = \frac{h^2}{2m} \left(\frac{3}{8\pi} n \right)^{2/3}. \quad (6.9)$$

At zero temperature, we can also easily compute other properties of the fermion gas. The internal energy is given by

$$U = \sum_r \epsilon_r \bar{n}_r \rightarrow \int_0^\infty d\epsilon f(\epsilon) \epsilon \bar{n}_{\text{FD}}(\epsilon), \quad (6.10)$$

and in the zero-temperature limit it approaches

$$U \xrightarrow{T \rightarrow 0} \int_0^{\epsilon_F} d\epsilon f(\epsilon) \epsilon = \frac{4\pi V}{h^3} (2m)^{3/2} \int_0^{\epsilon_F} d\epsilon \epsilon^{3/2}. \quad (6.11)$$

Comparing with Eq. (6.8), we can write this as

$$\begin{aligned} U &= \frac{3}{2} N \epsilon_F^{-3/2} \int_0^{\epsilon_F} d\epsilon \epsilon^{3/2} = \frac{3}{2} N \epsilon_F^{-3/2} \times \frac{2}{5} \epsilon_F^{5/2} = \frac{3}{5} N \epsilon_F \\ &= \frac{3h^2}{10m} \left(\frac{3}{8\pi} \right)^{2/3} \frac{N^{5/3}}{V^{2/3}}. \end{aligned} \quad (6.12)$$

This means that on average, each particle carries energy $(3/5)\epsilon_F$.

From Eq. (6.12), it is straightforward to calculate the pressure as

$$P = - \left(\frac{\partial U}{\partial V} \right)_{N,S} = \frac{2}{3} \times \frac{3h^2}{10m} \left(\frac{3}{8\pi} \right)^{2/3} \frac{N^{5/3}}{V^{5/3}} = \frac{2}{3} \frac{U}{V} = \frac{2}{5} \epsilon_F n, \quad (6.13)$$

where we could assume that S is constant because we are at zero temperature.

6.3 Ultrarelativistic Degenerate Electron Gas

There are also important examples of degenerate fermion gases in which the Fermi energy is much higher than the rest energy, $\epsilon_F \gg mc^2$, and therefore the fermions are **ultrarelativistic**. Using the ultrarelativistic density of translational states (3.67), and taking into account the two spin states, we find the full density of states

$$f(\epsilon) = g f_{\text{tr}}^{\text{UR}}(\epsilon) = \frac{8\pi V}{h^3 c^3} \epsilon^2. \quad (6.14)$$

We again calculate the number of particles N at zero temperature to determine the Fermi energy ϵ_F ,

$$N = \int_0^{\epsilon_F} f(\epsilon) d\epsilon = \frac{8\pi V}{h^3 c^3} \int_0^{\epsilon_F} \epsilon^2 d\epsilon = \frac{8\pi V}{3h^3 c^3} \epsilon_F^3. \quad (6.15)$$

Solving this for ϵ_F gives

$$\epsilon_F = hc \left(\frac{3}{8\pi} n \right)^{1/3}, \quad (6.16)$$

where $n = N/V$ is the number density of particles.

The internal energy at zero temperature is

$$U = \int_0^{\epsilon_F} \epsilon f(\epsilon) d\epsilon = \frac{8\pi V}{h^3 c^3} \int_0^{\epsilon_F} \epsilon^3 d\epsilon = \frac{3N}{\epsilon_F^3} \int_0^{\epsilon_F} \epsilon^3 d\epsilon = \frac{3N}{4} \epsilon_F = \frac{3hc}{4} \left(\frac{3}{8\pi} \right)^{1/3} \frac{N^{4/3}}{V^{1/3}}. \quad (6.17)$$

Because entropy is constant at zero temperature, we can easily calculate the pressure,

$$P = - \left(\frac{\partial U}{\partial V} \right)_{S,N} = \frac{1}{3} \frac{U}{V} = \frac{hc}{4} \left(\frac{3}{8\pi} \right)^{1/3} n^{4/3}. \quad (6.18)$$

6.4 White Dwarf Stars

White dwarfs are stars that have burned all their hydrogen. When that happens, the fusion reaction, which normally supports stars against gravitational forces, stops, and the star collapses. As we will see, when the density becomes high enough, the degeneracy pressure of the electrons in the star becomes strong enough to withstand the gravitational forces and the collapse is halted. The result star is a white dwarf.

We will do a simplified calculation, which will nevertheless give the correct picture of the physics and a reasonably good estimate of the quantities involved. We calculate the energy U_{tot} of the star, including the gravitational and internal components U_g and U_0 , and determine the radius of the star by minimising the total energy.

We assume that the density of the star is uniform, and that the gravitational potential energy is given by

$$U_g \approx - \frac{GM^2}{R}, \quad (6.19)$$

where G is Newton's constant, M is the mass of the star and R is its radius.⁷

The main contribution to the mass of the star comes from the protons and neutrons in it,

$$M \approx N_p m_p + N_n m_n, \quad (6.20)$$

⁷It is not difficult to calculate U_g exactly if the density is uniform, and the result has an extra factor of 3/5. However, it makes little sense to include this factor in the calculation, since the density of a real star would not be uniform, anyway.

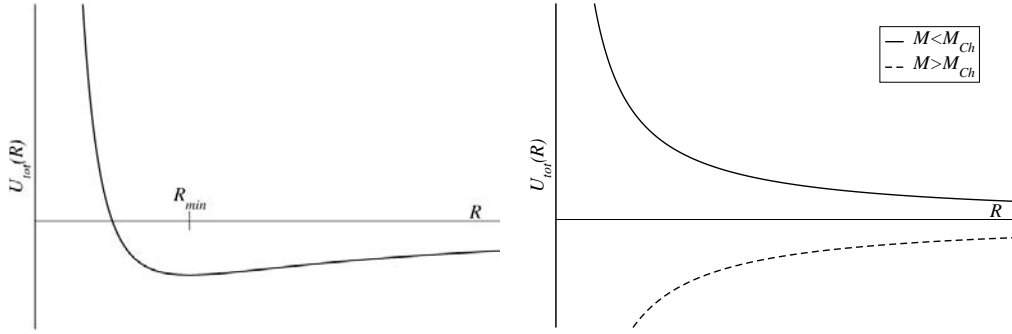


Figure 8: The total energy of a star as a function of its radius, for non-relativistic (left) and ultrarelativistic (right) electrons.

where N_p and N_n are the numbers of protons and neutrons, respectively, and m_p and m_n are their masses. In a typical atom, the number of neutrons is roughly equal to the number of protons, so we have $N_n \approx N_p$, and the masses of the two nucleons are also almost equal, $m_n \approx m_p$. Therefore, we can write the mass of the star as $M \approx 2N_p m_p$. Because a star is not expected to have a significant electric charge, it has to have an equal number of electrons and protons, $N_e = N_p$, and therefore

$$M \approx 2N_e m_p. \quad (6.21)$$

The internal energy depends on whether the electrons are non-relativistic or ultrarelativistic.

(i) If $\epsilon_F \ll m_e c^2$, the electrons are **non-relativistic**, and the internal energy is given by Eq. (6.12),

$$U_0 = \frac{3h^2}{10m_e} \left(\frac{3}{8\pi} \right)^{2/3} \frac{N_e^{5/3}}{V^{2/3}}, \quad (6.22)$$

where N_e is the number of electrons in the star and $V = (4\pi/3)R^3$ is its volume.

Using Eq. (6.21), we find

$$U_0 = \frac{3h^2}{20m_e m_p^{5/3}} \left(\frac{3}{8\pi} \right)^{4/3} \frac{M^{5/3}}{R^2} \equiv C_{\text{NR}} \frac{M^{5/3}}{R^2}, \quad (6.23)$$

where we have included all the constant factors into one constant C_{NR} .

The total energy is now

$$U_{\text{tot}}(R) = U_g(R) + U_0(R) \approx C_{\text{NR}} \frac{M^{5/3}}{R^2} - \frac{GM^2}{R}. \quad (6.24)$$

The shape of this function is shown in Fig. 8. It has a single minimum, whose location R_{min} we can find by setting the derivative to zero,

$$0 = \left. \frac{\partial U_{\text{tot}}}{\partial R} \right|_{R=R_{\text{min}}} = -2C_{\text{NR}} \frac{M^{5/3}}{R_{\text{min}}^3} + \frac{GM^2}{R_{\text{min}}^2}, \quad (6.25)$$

which gives

$$R_{\text{min}} = \frac{2C_{\text{NR}}}{G} M^{-1/3}. \quad (6.26)$$

This is the equilibrium radius of the star, at which the degeneracy pressure of the electrons balances the gravitational force. Interestingly, the more massive the star is, the smaller its radius.

The Fermi energy can be calculated using Eq. (6.9), and is proportional to

$$\epsilon_F \propto n^{2/3} \propto \left(\frac{M}{R_{\min}^3} \right)^{2/3} \propto (M^2)^{2/3} = M^{4/3}. \quad (6.27)$$

This shows that the more massive the star is, the higher the Fermi energy. If the star is massive enough, the Fermi energy will exceed $m_e c^2$, and the assumption that the electrons are non-relativistic breaks down.

- (ii) If $\epsilon_F \gg m_e c^2$, the electrons are **ultrarelativistic** and we can use the results from Section 6.3. According to Eq. (6.17), the internal energy is

$$U_0 = \frac{3hc}{4} \left(\frac{3}{8\pi} \right)^{1/3} \frac{N_e^{4/3}}{V^{1/3}} = \frac{3hc}{8m_p^{4/3}} \left(\frac{3}{8\pi} \right)^{2/3} \frac{M^{4/3}}{R} \equiv C_{\text{UR}} \frac{M^{4/3}}{R}. \quad (6.28)$$

The total energy is therefore

$$U_{\text{tot}}(R) = U_g(R) + U_0(R) \approx C_{\text{UR}} \frac{M^{4/3}}{R} - \frac{GM^2}{R} = \left(C_{\text{UR}} M^{4/3} - GM^2 \right) \frac{1}{R}. \quad (6.29)$$

This function has no minimum, so there is no equilibrium state. Depending on the sign of the factor in the brackets, we can have two different situations:

If $M < (C_{\text{UR}}/G)^{3/2}$, the factor is positive (see Fig. 8), and the star expands until the Fermi energy has fallen close to m_e . Then the ultrarelativistic approximation is not valid any more, and non-relativistic effects stabilise the star.

If $M > (C_{\text{UR}}/G)^{3/2}$, the factor is negative, and there is nothing to prevent the star from shrinking. The star collapses until the degeneracy pressure of neutrons becomes strong enough to stop it, and it becomes a neutron star. If the star is massive enough even the neutrons cannot withstand the gravitational force, and it becomes a black hole.

The limiting mass

$$M_{\text{Ch}} = \left(\frac{C_{\text{UR}}}{G} \right)^{3/2} \quad (6.30)$$

is known as the **Chandrasekhar mass**, and it is the highest possible mass that a stable white dwarf can have. Its value is approximately 1.4 solar masses. Any star that is less massive than this, for instance our Sun, will form a white dwarf when it runs out of hydrogen. More massive stars become neutron stars or black holes.

A Mathematical Results

A.1 Combinatorics

How many ways are there to pick M objects out of the total of N ?

- (a) *Permutations with Repetition*: If each object can be picked several times, and the order in which the objects were picked matters, the number of possible ways is

$$N^M. \quad (\text{A.1})$$

- (b) *Permutations without Repetition*: If each object can be picked only once, and the order in which the objects were picked matters, the number of possible ways is

$$N(N-1)\cdots(N-M+1) = \frac{N!}{(N-M)!} \quad (\text{A.2})$$

- (c) *Combinations with Repetition*: If each object can be picked several times, and the order in which the objects were picked does not matter, the number of possible ways is

$$\frac{(N+M-1)!}{M!(N-1)!}. \quad (\text{A.3})$$

- (d) *Combinations without Repetition*: If each object can be picked only once, and the order in which the objects were picked does not matter, the number of possible ways is

$$\frac{N!}{M!(N-M)!}. \quad (\text{A.4})$$

A.2 Stirling's Formula

When $N \gg 1$,

$$\ln N! = N \ln N - N + \frac{1}{2} \ln 2\pi N + O(1/N) \approx N \ln N - N. \quad (\text{A.5})$$

Simple derivation of the less accurate form: Let us write

$$\ln N! = \sum_{p=1}^N \ln p, \quad (\text{A.6})$$

and approximate this sum by an integral,

$$\ln N! \approx \int_1^N dp \ln p = N \ln N - N + 1 \approx N \ln N - N. \quad (\text{A.7})$$

A.3 Gaussian Integrals

Gaussian integrals are of the form

$$I_n(a) = \int_0^\infty dx x^n e^{-ax^2}, \quad (\text{A.8})$$

with $a > 0$. The general result for them is

$$I_n(a) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{2a^{(n+1)/2}}, \quad (\text{A.9})$$

where $\Gamma(x)$ is the Euler gamma function. For integer arguments, the gamma function becomes just the factorial,

$$\Gamma(k+1) = k!. \quad (\text{A.10})$$

If n is odd (i.e., we can write $n = 2k + 1$), the result can therefore be written in a simple form

$$I_{2k+1}(a) = \frac{k!}{2a^{k+1}}. \quad (\text{A.11})$$

For a few of the lowest even values of n , one has

$$I_0(a) = \frac{1}{2} \sqrt{\frac{\pi}{a}}, \quad (\text{A.12})$$

$$I_2(a) = \frac{1}{4a} \sqrt{\frac{\pi}{a}}, \quad (\text{A.13})$$

and

$$I_4(a) = \frac{3}{8a^2} \sqrt{\frac{\pi}{a}}. \quad (\text{A.14})$$

By a change of variables, one find

$$\int_0^\infty dt t^{x-1} e^{-\beta t} = \frac{\Gamma(x)}{\beta^x} \quad (\text{A.15})$$

for $x > 0$. For example,

$$\int_0^\infty dt t^{1/2} e^{-\beta t} = \frac{\Gamma(3/2)}{\beta^{3/2}} = \frac{1}{2} \sqrt{\frac{\pi}{\beta^3}}. \quad (\text{A.16})$$

A.4 Other Useful Integrals

Another common integral is

$$\int_0^\infty \frac{x^a}{e^x - 1} = \Gamma(1+a)\zeta(1+a), \quad (\text{A.17})$$

where $a > 0$ and $\zeta(x)$ is the Riemann zeta function. Useful examples of this are

$$\int_0^\infty dx \frac{x^3}{e^x - 1} = \Gamma(4)\zeta(4) = \frac{\pi^4}{15}, \quad (\text{A.18})$$

and

$$\int_0^\infty dx \frac{x^{1/2}}{e^x - 1} = \frac{\sqrt{\pi}}{2} \zeta\left(\frac{3}{2}\right) \approx 2.315. \quad (\text{A.19})$$

A related integral is

$$\int_0^\infty dx x^2 \ln(1 - e^{-x}) = -2\zeta(4) = -\frac{\pi^4}{45}. \quad (\text{A.20})$$

A.5 Sums

A.5.1 Geometric Series

The sum of a geometric series can be calculated exactly,

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad (\text{A.21})$$

assuming that $x < 1$.

A.5.2 Multiple Summation

Manipulating multiple sums is a critical skill in statistical physics. We frequently encounter sums like

$$\sum_{r_1} \cdots \sum_{r_N} f(r_1, \dots, r_N), \quad (\text{A.22})$$

where f is some function of the summation indices r_i .

The notation \cdots generally means repeating the same symbol or expression. When the dots are between summations, it means multiple summation, i.e., summing over all N indices r_1, \dots, r_N . In general, this is not a product. For instance,

$$\sum_{r_1} \cdots \sum_{r_4} f(r_1, \dots, r_4) \equiv \sum_{r_1} \sum_{r_2} \sum_{r_3} \sum_{r_4} f(r_1, r_2, r_3, r_4). \quad (\text{A.23})$$

What multiple summation means in practice is that you first do the innermost sum (r_N). You get a result that depends on the indices r_1, r_2, \dots, r_{N-1} . You sum this result over r_{N-1} , and the result of that sum over r_{N-2} and so on, until you have summed over all the indices r_i .

Often in statistical physics, the summed function $f(r_1, \dots, r_N)$ is an exponential of the form

$$f(r_1, \dots, r_N) = \exp \left[\sum_{i=1}^N x_i(r_i) \right], \quad (\text{A.24})$$

where each x_i depends only on one of the indices r_i . Remember that the exponential of a sum is a product of exponentials, so we have

$$\exp \left[\sum_{i=1}^N x_i(r_i) \right] = \prod_{i=1}^N e^{x_i(r_i)} \equiv e^{x_1(r_1)} \cdots e^{x_N(r_N)}. \quad (\text{A.25})$$

In this case, the multiple sum factorises because from the point of view of the sum over r_i any factor $e^{x_j(r_j)}$ with $j < i$ is a constant and can be taken outside the sum,

$$\begin{aligned} \sum_{r_1} \cdots \sum_{r_N} \exp \left[\sum_{i=1}^N x_i(r_i) \right] &= \sum_{r_1} \cdots \sum_{r_N} e^{x_1(r_1)} \cdots e^{x_N(r_N)} = \sum_{r_1} e^{x_1(r_1)} \cdots \sum_{r_N} e^{x_N(r_N)} \\ &= \left(\sum_{r_1} e^{x_1(r_1)} \right) \cdots \left(\sum_{r_N} e^{x_N(r_N)} \right) \equiv \prod_{i=1}^N \left(\sum_{r_i} e^{x_i(r_i)} \right) \end{aligned} \quad (\text{A.26})$$

Furthermore, if all the functions $x_i(r)$ are the same, i.e., $x_i(r) = x(r)$, then all the factors in this product are equal, and we have

$$\sum_{r_1} \cdots \sum_{r_N} \exp \left[\sum_{i=1}^N x(r_i) \right] = \left(\sum_r e^{x(r)} \right)^N. \quad (\text{A.27})$$

B General Expression for Entropy

Consider first a system consisting of just a single particle, with single-particle states $r \geq 0$. Its *macrostate* is described by some ensemble, and we denote the probability that the particle is in state r by p_r . Our aim is to calculate the entropy of this macrostate.

To do this, it turns out to be useful to consider a system of N noninteracting distinguishable particles first. Even though the particles are distinguishable, we can still consider the occupation number n_r of state r , defined as the number of particles in state r (see Section 2.2). However, knowing the occupation numbers n_r of all the single-particle states r . This does *not* specify the microstate of the whole system, because the particles are distinguishable and we do not know which particle is in which state. Instead, the occupation numbers define a *macrostate* of the system. The corresponding *microstates* are different permutations of the states of the particles.

As the first step, let us calculate the entropy S_N of this macrostate. Following the fundamental postulate, we assume that all of the microstates (i.e. permutations of single-particle states) have equal probability. According to Eq. (2.18), the entropy is given by the multiplicity Ω of the macrostate, i.e., how many microstates it corresponds to. The microstates are labelled by N integers (r_1, \dots, r_N) , consisting of n_0 zeros, n_1 ones, n_2 twos etc., and to obtain the multiplicity we need to count the number of different permutations.

If all the integers were different, the number of permutations would be $N!$. However, the $n_1!$ permutations that swap two 1s, do not change the microstate, and neither do the $n_2!$ permutations that swap two 2s, not the $n_r!$ permutations that swap two r s. Thus, the number of *different* permutations is (as a fairly straightforward generalisation of Eq. (A.4))

$$\Omega = \frac{N!}{n_1!n_2!\dots} = \frac{N!}{\prod_{r=0}^{\infty} n_r!}. \quad (\text{B.1})$$

The entropy of this macrostate is therefore

$$S_N = k_B \ln \Omega = k_B \left(\ln N! - \sum_{r=0}^{\infty} \ln n_r! \right), \quad (\text{B.2})$$

and using Stirling's formula (A.7), we find

$$S_N = k_B \left(N \ln N - N - \sum_{r=0}^{\infty} n_r \ln n_r + \sum_{r=0}^{\infty} n_r \right) = -k_B \sum_{r=0}^{\infty} n_r \ln \frac{n_r}{N}, \quad (\text{B.3})$$

where we used the fact that

$$\sum_{r=0}^{\infty} n_r = N. \quad (\text{B.4})$$

As the second step, let us go back to the original setup in which the particles follow the probability distribution p_r , rather than having fixed occupation numbers n_r . The mean occupation number is given simply by

$$\bar{n}_r = N p_r. \quad (\text{B.5})$$

When N is very large, we can replace n_r with \bar{n}_r in Eq. (B.3). Therefore we have for a system of $N \gg 1$ particles

$$S_N = -k_B \sum_{r=0}^{\infty} \bar{n}_r \ln \frac{\bar{n}_r}{N} = -N k_B \sum_{r=0}^{\infty} p_r \ln p_r. \quad (\text{B.6})$$

Assuming that entropy is extensive, the entropy of a single particle is then

$$S = \frac{S_N}{N} = -k_B \sum_{r=0}^{\infty} p_r \ln p_r. \quad (\text{B.7})$$

As the final step, we note that this result is general, and does not depend on what our “particle” is. Therefore, even if we consider some macroscopic system in some ensemble p_α , we can think of it as the “particle”. The microstates α of our macroscopic system would correspond to the “single-particle states” r . Applying Eq. (B.7), we see that for any system, even a macroscopic one, we have

$$S = -k_B \sum_{\alpha} p_{\alpha} \ln p_{\alpha}. \quad (\text{B.8})$$

This is the general result for the entropy of an ensemble. It is easy to check that it agrees with the Boltzmann expression (2.18) in the case of the microcanonical ensemble.

C Notation

C.1 Indices

- α Microstate
- i Particle
- r Single-particle state

C.2 Quantities

- β Inverse temperature
- ϵ_r Energy of single-particle state r
- Φ Grand potential
- μ Chemical potential
- Ω Multiplicity
- E Total energy
- f Density of states
- F Free energy
- g Degeneracy
- n_r Occupation number of single-particle state r
- N Number of particles (number density= n)
- p_α Probability of microstate α
- \vec{p} Momentum
- P Pressure
- S Entropy (entropy density= s)
- T Temperature
- U Internal energy (energy density= ρ)
- V Volume
- \mathcal{Z} Grand partition function
- Z Partition function

C.3 Constants

- c Speed of light (= 299792458m/s)
- h Planck's constant ($\approx 6.626 \times 10^{-34} \text{m}^2\text{kg/s}$ and $\hbar = h/2\pi$)
- k_B Boltzmann's constant ($\approx 1.381 \times 10^{-23} \text{m}^2\text{kg/s}^2\text{K}$)

Index

- Black-body radiation, 27
- Boltzmann distribution, 14
- Boltzmann entropy, 10
- Bose-Einstein condensation, 32
- Bose-Einstein distribution, 26
- Bosons, 7

- canonical ensemble, 15
- Chandrasekhar mass, 38
- chemical potential, 23
- classical equipartition of energy, 20

- degeneracy, 18
- degenerate, 18, 34
- density of states, 18
- distribution, 25

- ensemble, 3
- equation of state, 3
- equilibrium, 4
- ergodic hypothesis, 12

- Fermi Energy, 34
- Fermi temperature, 34
- Fermi-Dirac distribution, 26
- Fermions, 8
- free energy, 16
- fundamental postulate, 4

- grand canonical ensemble, 23
- grand partition function, 23
- grand potential, 23

- isolated system, 4

- laws of thermodynamics, 3

- macroscopic, 2
- macroscopic variables, 3
- macrostate, 3
- microcanonical ensemble, 4
- microscopic, 2
- microstate, 2
- multiplicity, 5

- non-relativistic, 20, 37

- occupation number, 6

- Pauli exclusion principle, 8
- Planck radiation law, 28

- single-particle states, 5, 18
- spectral energy density, 28
- Stefan-Boltzmann constant, 29
- Stefan-Boltzmann law, 29

- temperature, 10
- thermodynamic limit, 16

- ultrarelativistic, 20, 36