

Statistical Mechanics

TanXG lecture notes on theoretical physics

D.E.Liang
tanxg2023@outlook.com



Statistical Mechanics

Introduction

LD

Contents

1 Foundations of Statistical Mechanics	1
1.1 Basic Concepts for a System	2
1.1.1 What's the objects that we are interested in?	2
1.1.2 Equilibrium	2
1.1.3 Macrostates and Macroscopic Parameters	3
1.1.4 Microstates	4
1.2 Postulate of Statistical Mechanics	5
1.2.1 Postulate of Statistical Mechanics	5
1.2.2 A simple example — quantum particles in a 3D box	5
1.3 Irreversibility and Boltzmann Entropy	7
1.3.1 Irreversibility	7
1.3.2 Boltzmann entropy	8
1.4 Ensembles in Statistical Mechanics	11
1.4.1 Statistics with ensembles	11
1.4.2 Microcanonical ensembles	12
1.4.3 Canonical Ensembles	13
1.4.4 Grand Canonical Ensembles	17
1.4.5 Gibbs Ensembles and Pressure	19
1.4.6 How to Understand the Names of Different Ensembles?	19
1.5 Partition Function	21
1.5.1 Partition Function	21
1.5.2 Simple Applications of Partition Function	21
2 Ideal Classical Gases	31
2.1 Classical Partition Function	32
2.1.1 Classical Partition Function	32
2.1.2 Monatomic Molecules	34
2.1.3 Polyatomic Molecules	35

CONTENTS

2.2	Gibbs Paradox and Entropy	39
2.2.1	Gibbs Paradox	39
2.2.2	Gibbs Modification for partition functions	40
2.2.3	Entropy	40
2.3	Pressure and Energy	42
2.3.1	Pressure	42
2.3.2	Energy	43
2.4	Maxwell Distribution	45
2.4.1	Maxwell Distribution of Velocity	45
2.4.2	Maxwell Distribution of Speed	45
3	Quantum Statistics I	47
3.1	Energy of Ideal Gases Revisited	48
3.1.1	Translation	48
3.1.2	Rotation	49
3.1.3	Vibration	50
3.2	Photon Gases	52
3.2.1	Density of States in 3D box	52
3.2.2	Plank Distribution and Blackbody Radiation	54
3.2.3	Radiation Pressure	59
3.3	Heat Capacity of Solid	61
3.3.1	Dulong - Petit Law	61
3.3.2	Einstein's Model	62
3.3.3	Debye's Model	63
4	Quantum Statistics II	67
4.1	Distribution of Particle number	68
4.1.1	Maxwell-Boltzmann Distribution	68
4.1.2	Bose-Einstein Distribution	68
4.1.3	Fermi-Dirac Distribution	69
A	Gauss Integrals	71
B	Stirling Formula	73
C	我与热力学二三事	75
C.1	入门教程	76
C.2	新手村	76
C.3	打副本	78

CONTENTS

C.4 解锁天赋技能	79
----------------------	----

CONTENTS

Chapter 1

Foundations of Statistical Mechanics

1.1 Basic Concepts for a System

1.1.1 What's the objects that we are interested in?

In a word, the objects that we are interested in are mainly the *systems with a large number of constituents*.

In the many places of physics, we only need to deal with single object —one particle, one planet, one charge, one electron moving around one nucleus. However, there also exists other kind of systems which consist of a large number of constituents¹. In some situation, it can be important to figure out the physical properties of such systems, thus asking us to find efficient methods to research objects with large number of constituents.

It turns out to be impossible, and also unnecessary, to get knowledge of the whole system by studying its constituents detailly one by one. Instead, what we really need to care about are the concepts and methods that only hold for the entirety, such as entropy, partition function and so on. We'll meet them as we proceed.

The entirety is essential, that's what we should always keep in mind when taking the trip of statistical mechanics.

Finally, let's classify different kinds of systems. *Isolated systems* don't have any interactions with other systems or environment; *closed systems* can have energy exchanges with other systems or environment but no matter exchanges; and *open systems* can have both energy and matter exchanges with other systems or environment.

1.1.2 Equilibrium

Daily experience tells us that, as time goes on, a system will ultimately reach a kind of state that all of its observable macroscopic properties become uniform within the system and don't change with time anymore. We say this kind of state *equilibrium*.

Strictly speaking, it is meaningless to define the macroscopic physical quantities such as pressure or temperature for the whole system if it's not in equilibrium. Because the properties are not yet uniform within the system in the state of non equilibrium. In this non equilibrium situation we could only define the local pressure or local temperature in the small regions within the system, and of course different regions in the system maybe have different yet locally uniform magnitudes of such quantities.

¹For example, the simplest system like that is gas. In our house, 1L air can has 10^{22} molecules in it.

There's also another kind of occasion for the system that all of the macroscopic quantities are unchanged over time but not uniform within the system. However, the system like this can't be defined as in equilibrium, instead we ought to call it in *stable state*².

We can define lots of macroscopic parameters for a system as long as it's in equilibrium, which will be discussed next subsection.

Statistical mechanics can be divided into equilibrium statistical mechanics and non equilibrium statistical mechanics. In our lecture notes, we will only focus on the equilibrium statistical mechanics. Which means we will only care about the systems in equilibrium.

1.1.3 Macrostates and Macroscopic Parameters

Facing with a system in equilibrium, we can ask a series of questions like kids : How hot is it? What colour is it? Is it big or small? Actually, what we've mentioned in these questions are concerned with the *macroscopic parameters* of the system. Some of the typical macroscopic parameters are pressure p , absolute temperature T , volume V . These macroscopic parameters describe the macroscopic properties of the system.

For a simple system in equilibrium, we can confirm its physical properties only with p, V and T . But in other occasions with more complexities, we'll need more macroscopic parameters. For example, if there exists electromagnetic field, we still need to confirm the polarization intensity \mathbf{P} and the magnetization intensity \mathbf{M} ; else if we mix different kinds of gases together, their percentages and total number of constituents N ought to be taken into account³.

In our lecture notes, we won't consider the parameters \mathbf{P} and \mathbf{M} .

Then comes to the term *macrostate*, which isn't difficult to understand literally. We can identify a macrostate of the system by identifying all of its macroscopic parameters⁴.

In an equilibrium system, some of the macroscopic parameters are not independent with each other, instead, they may be contained in one equation. This kind of equation is essential in thermodynamics, and are called *equation of state*. There's

²One of simple examples for this kind of situation is that you put a metal bar between a hot body and cold body with consistent temperatures. Then, along the metal bar, different positions will have different, and stable, temperatures. We can't say this metal bar is in equilibrium because there exists heat flow within it. It's in stable state instead.

³We need to consider N because there could be chemical reactions in the system which may change the magnitude of N .

⁴As a result, we also would like to call macroscopic parameters the *state parameters* of the system.

no way except for doing experiments to figure out the equation of state of a system in thermodynamics. But we can deduce the equation of state by using the methods of statistical mechanics theoretically, which is one of the reasons why the theory of statistical mechanics is so fascinating.

Besides, there exist lots of physical quantities in thermodynamics, which are functions of state parameters. These quantities, such as energy E and thermodynamic potentials, are called *functions of state*.

1.1.4 Microstates

Microstate is the dual concept of Macrostate.

Simply speaking, to identify a microstate of a system, we must identify all the concrete microstate of each constituent of the whole system⁵, which can be practically impossible.

Nonetheless, fortunately, what we really need is the number of microstates, instead of the mathematical formalism. For a determined macrostate, usually it corresponds to lots of different microstates, and we denote Ω as the number of microstates of a determined macrostate. It's not hard to understand that Ω oughts to be a function of macroscopic parameters, or even be a function of functions of macrostate⁶.

In fact, there are lots of examples which reveal that Ω can be a ridiculously large number⁷, and no one can manage to figure out the exact magnitude of it even for the most simple systems, like a bottle of air. Nonetheless, fortunately again, ultimately we'll find that, on the one hand, we don't need the mathematical form of Ω but only need to know that Ω is large; on the other hand, although in some context we have to use the concretely form of Ω , some useful approximation may help.

⁵This way of comprehending can be fine in most cases, but can be wrong in some situations such as when we are talking about the photon gas. We will see this in chapter 3.

⁶Because Ω is actually a function of macrostate.

⁷For example, see the simple example I provide in next section.

1.2 Postulate of Statistical Mechanics

1.2.1 Postulate of Statistical Mechanics

The postulate of statistical mechanics is not hard to understand, yet it can be a very useful as well as powerful principle for us. It will turn out that the essence why the methods of statistical mechanics can deduce the same results as thermodynamics or experiments is that we believe in the postulate of it.

For an isolated system in equilibrium, firstly there suppose to be some *constraints* on it, such as the total energy E_{total} or the total number of constituents N are fixed. Then, the postulate of statistical mechanics tells us that *for such an isolated⁸ system in equilibrium with certain constraints, all of its achievable microstates are of equal probability.*

1.2.2 A simple example —quantum particles in a 3D box

To illustrate what's the meaning of the postulate, let's consider a simple example. Suppose we have a box with volume of $V = L^3$, and 2 identical but distinguishable quantum particles are flying around in this box. According to quantum mechanics, their energy levels are

$$E = (n_x^2 + n_y^2 + n_z^2) \frac{\pi^2 \hbar^2}{2mL^2} \stackrel{\text{def}}{=} (n_x^2 + n_y^2 + n_z^2) E_0$$

where n_x , n_y and n_z are all positive integers. So the microstates of the system is actually represented by 6 quantum numbers $(n_{1x}, n_{1y}, n_{1z}; n_{2x}, n_{2y}, n_{2z})$. Suppose that the constraint for this system is $E_{total} = E_1 + E_2 = 9E_0$. Then the achievable microstates for the system are

$$(2, 1, 1; 1, 1, 1)$$

$$(1, 2, 1; 1, 1, 1)$$

$$(1, 1, 2; 1, 1, 1)$$

$$(1, 1, 1; 2, 1, 1)$$

$$(1, 1, 1; 1, 2, 1)$$

$$(1, 1, 1; 1, 1, 2)$$

The postulate tells us that if we "have a look" at this system, the probability that we find the system in any one of these 6 states are same, which is $1/6$. And here $E_{total} =$

⁸Here the term "isolated" must be satisfied since for an non isolated system, different microstates may have different probability to be found. Generally that's to say there oughts to be a probability distribution for the microstates, which is to be discussed next section.

$9E_0$ can be viewed as a macrostate, while the 9 states above are its microstates. Remember what we say about Ω ? If we view Ω as a function of E_{total} here, then we can get that $\Omega(9E_0) = 6$.

All of the results above are simple and kind of obvious, because we only have 2 particles in the isolated box. As we add more particles in it, the circumstance will get more and more complex rapidly. There is indeed another example, with 5 particles in the box, which I met as a mini project when taking the course of *thermodynamics and statistical mechanics*. At that time we were required to find some answers by coding with Python, and below are some interesting results.

If there are 5 identical but distinguishable particles in the isolated box with $E_{\text{total}} = 50E_0$ as a constraint, then the total number of achievable microstates are 197940, i.e. $\Omega(50E_0) = 197940$. However, the number of microstates with the 1st particle having energy of $38E_0$ ⁹ is only 9. That's to say, when the system reaches equilibrium, you only have the probability of $\frac{9}{197940} = 0.000045$ to find the situation that almost all the energy are spontaneously gathered in the first particle. This kind of tiny probability just means IMPOSSIBLE. This, is a simple demonstration to the irreversibility, which is detailedly discussed next section.

⁹ $38E_0$ is the largest energy that one particle can have in this context, isn't it?

1.3 Irreversibility and Boltzmann Entropy

1.3.1 Irreversibility

The postulate of statistical mechanics embodies the *irreversibility* of some natural process, which is directly related to the 2nd law of thermodynamics.

Again let's think about the scenario similar to what we mentioned last section. Suppose there are 5 particles put in an isolated 3D box. Here, 1 particle with energy of $38E_0$ is in the left part of the box, while 4 particles each of which has energy of $3E_0$ are in the right part of the box. There is a partition in the middle of the box. Then we pull out the partition, which can be viewed as changing the constraint. Subsequently, the system will evolve with time isolatedly. In the end, it will get to equilibrium. According to the result showed in last subsection, we can figure out that after the system getting to the equilibrium, the probability we find the system go back to the origin condition is 0.000045. In other words, it is impossible for the system to go back, which reveals a kind of irreversibility.

The postulate of statistical mechanics tells us the direction of the time evolution for an isolated system after changing its constraint. The statement that I prefer most about this is shown below

After changing the constraint, whatever the initial conditions are, the system will ultimately approach¹⁰ macrostates that have most number of microstates under new constraint as equilibrium states.

In next subsection, we will argue that this statement provides an opportunity to bridge the concept of entropy for isolated system with the number of its achievable microstates.

Because of the low probability, it's impossible to find an equilibrium system that fluctuates too far away from the certain macrostates mentioned above. But tiny fluctuations among similar macrostates are allowed, accompanied by tiny fluctuations in observable magnitudes which can be finely ignored in macroscopic scale¹¹.

Last but not least, what I also want to emphasize is that the postulate of statistical mechanics doesn't give us any knowledge of the process of evolving. We don't know the dynamics describing how the systems transfer from one state to another state in these lecture notes. We only know how to deal with equilibrium states. So we conclude the irreversibility from the result, instead of from the evolving process.

¹⁰Or nearly approach. I state like this because sometimes there are many "similar" macrostates all have large number of microstates and are closed to each other, hence the equilibrium system might jump among these kinds of macrostates.

¹¹This also accounts for why in equilibrium the system can show stable.

1.3.2 Boltzmann entropy

We've met the concept of entropy in the courses of thermodynamics. Denoted as S , generally speaking the entropy for simple system is a state function of energy E , volume V and number of constituents N

$$S = S(E, V, N)$$

We have also known from thermodynamics that the key property of S is that **for an isolated system, its entropy never decrease. If the isolated system undergoes a reversible process, then $\Delta S = 0$; while if it undergoes an irreversible process, then $\Delta S > 0$.**

We can interpret the term "process" as changing the constraint and subsequently the system transfer from one equilibrium state to another equilibrium state. Thus, for the new equilibrium state under an irreversible process, its entropy will increase. And it's impossible for the system to go back to the original state since that's what the "irreversible" means.

This kind manner is very similar to what we said last section about the irreversibility. After changing the constraint, the increase of Ω for the new equilibrium state leads to the impossibility for the system to return the original state. So in this sense, the increase of thermodynamic entropy is equivalent to the increase of Ω — both of the increase represent the irreversibility.

A Helpful Scenario

There is a helpful scenario to help you understand the arguments above.

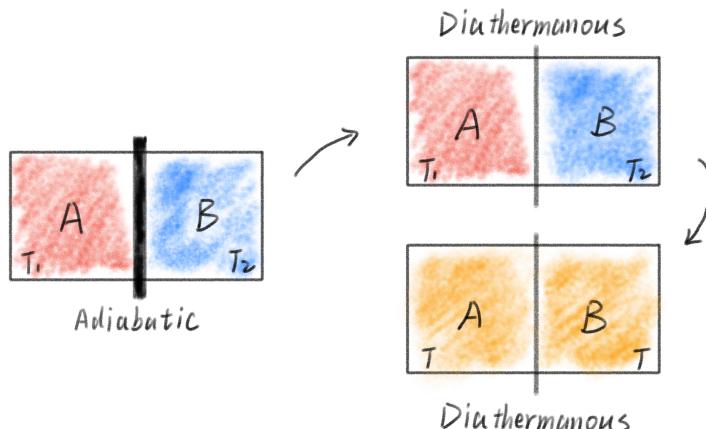


Figure 1.1: A helpful scenario

Imagine an isolated box divided by a partition. At beginning, the partition is adiabatic, and the gas in part A has higher temperature T_1 , while the gas in part B has lower temperature T_2 . The whole system is in equilibrium, with the number of achievable microstates $\Omega_{\text{total}} = \Omega_A \Omega_B$.

Then, we change the adiabatic partition into a diathermanous one, which should be viewed as the changing of constraint. Several minutes later, the system will reach a new equilibrium state with equal temperature T for A and B . And there's no denying that the new number of achievable microstates Ω'_{total} satisfy that

$$\Omega'_{\text{total}} \gg \Omega_{\text{total}} \quad (1.1)$$

(1.1) holds because the initial macrostate is just ONE of the achievable macrostate for the final equilibrium state.

Then, on the one hand, thermodynamics tells us that the entropy has increased for the system, and the process is irreversible¹²; on the other hand, postulate of statistical mechanics tells us that the system can't go back spontaneously because $\Omega' \gg \Omega$ and thus the probability of this happening spontaneously is $\frac{\Omega}{\Omega'} \ll 1$, which simply means impossible.

The Formula of Boltzmann Entropy

All the arguments above implies that there oughts to be a connection between Ω and the thermodynamic entropy for an isolated system. Our next goal is to find it.

Suppose there is a function between these two quantities, and we generally denote it as

$$S = f(\Omega)$$

again let's use the scenario above. Considering that the entropy is additive, thus before changing the constraint we have

$$S_{\text{total}} = S_A + S_B$$

Hence we have

$$f(\Omega_A \Omega_B) = f(\Omega_A) + f(\Omega_B)$$

Thus there oughts to have a log function. We define the entropy for an isolated system as

$$S = k \ln \Omega \quad (1.2)$$

¹²Remember the Clausius' statement of 2nd law of thermodynamics.

where k should be a proportionality coefficient. This kind of form of entropy is called *Boltzmann entropy*.

You should keep in mind that the Boltzmann entropy is for isolated system under certain constraints. For other kinds of systems, how to understand their entropy? We will argue this latter.



Figure 1.2: $S = k \log W$ on Boltzmann's tombstone

In fact, the coefficient k in (1.2) is not just a proportionality constant but exactly a meaningful natural constant named *Boltzmann constant*, thus make this formula more fascinating. We've met the Boltzmann constant in the courses of thermodynamics¹³. Is it a coincidence that we also have k in the formula of Boltzmann entropy? Well, I have to say, the answer is no. Theoretically speaking the truth is that it is (1.2) that the principal origin of k , instead of other methods to deduce k .

However, historically speaking, the formula (1.2) was first appear in Max Plank's paper instead of Boltzmann themselves. But people still engraved this formula on Boltzmann's tombstone in memory of this great yet unfortunate statistical physicist.

¹³Perhaps in the formula $p = nkT$ for the first time.

1.4 Ensembles in Statistical Mechanics

1.4.1 Statistics with ensembles

What and Why are Ensembles?

Generally speaking, for a system in equilibrium under certain constraints, it may have lots of different macrostates, and each one of the macrostates may contain lots of different achievable microstates. The equilibrium system is jumping among these microstates with some certain *distribution of probability*, thus all the results under measurements have a sense of average. So that's why the theory of statistics matters a lot in dealing with our interested systems, because initially what we have are probability distributions before addressing any practical problems. We are dealing with probability distribution, and using the probability distribution to calculate the average of other quantities.

Thus, there's no reason for us to think that this system is in a certain microstate. Instead, we should view any equilibrium system as in an *ensemble*. An ensemble of a system is an imaginary set of a large number of identical systems in equilibrium. At a certain time, different imaginary systems in the same ensemble will be in different microstates.

Statistics with Ensembles

The notion of ensembles provides us a good pointview to statistics about the systems.

Firstly, let's talk about probability. What I argue below are practical definitions by using the concept of ensemble.

Suppose that O_i is the magnitude of a macroscopic observable for the system in the i -th macrostate. If we do imaginary measurements on each system in the ensemble, we will get a certain macrostate for each one of the system. Suppose that we totally have N systems in the ensemble, and under our measurements, we find that N_i systems have the result of O_i . Thus the frequency that we get O_i as a result is

$$f_i = \frac{N_i}{N} \tag{1.3}$$

According to the *Law of large number* in statistics, as $N \rightarrow \infty$, the frequency f_i equals to the probability p_i that the system stays in the macrostate of O_i .

Furthermore, we denote $\Omega(O_i)$ as the number of microstates of the macrostates of O_i . Then it's obvious that the probability that the system stays in the corresponding

microstate $|i\rangle$ equals to

$$p_i = \lim_{N \rightarrow \infty} \frac{N_i / \Omega(O_i)}{N} \quad (1.4)$$

What's more, although for the sake of clarity, we only have one macroscopic parameter in (1.4), there can be, and ought to be, more parameters to indentify a macrostate, where we should take all of these parameters into account in the formula of probability¹⁴.

Secondly, armed with probability, we can calculate the average and dispersion of O . In our context of ensemble, the average of O is

$$\langle O \rangle = \lim_{N \rightarrow \infty} \frac{\sum_i N_i O_i}{N} = \sum_i p_i O_i \quad (1.5)$$

which is called the *ensemble average*. In contrast to the *time average*, the ensemble average is the average over all the systems in the same ensemble at the same time. From now on, all the averages appear in our lecture notes will be the ensemble averages, unless we make special illustration.

And we can also know the dispersion is

$$\sigma_O^2 = \langle O^2 \rangle - \langle O \rangle^2 \quad (1.6)$$

From now on, all the average

All the things above are just a brief review of the basic knowledge of statistics. Actually, in my point of view, the idea of ensembles is nothing but the familiar *independent and repeated experiments* that we learnt in statistics.

1.4.2 Microcanonical ensembles

Ensembles of isolated systems with fixed E , V and N are called *microcanonical ensembles*.

If we use the macroscopic quantities (E, V, N) to represent a macrostate of a system, then obviously all the systems in microcanonical ensemble are in the same macrostate but can be in different microstates. Thus, it is evident that, in microcanonical ensemble, the probability to find the system in the macrostate E is 1; while, according to (1.4) and the postulate of statistical mechanics, the probability to find

¹⁴For example, just look at the formula (1.7) below.

the system in any one of the microstates $|n\rangle$ is

$$p_n = \lim_{N \rightarrow \infty} \frac{N/\Omega(E, V, N)}{N} = \frac{1}{\Omega(E, V, N)} \quad (1.7)$$

The formula (1.7) may seem to be too obvious, however, it turns out to be a basis for us to proceed.

1.4.3 Canonical Ensembles

The microcanonical ensembles describe the isolated systems, but it's not the most common case in nature. Most of the systems that we interested in are not isolated, in contrast they are consistently interacting with the environment. Thus, the E , V and N of the systems may not be constants anymore. Instead, the quantities like temperature T , pressure p and chemical potential μ can be constants.

Now, let's consider systems with fixed N and V , yet can have thermal interaction with the environment¹⁵. The ensembles of such kind of systems in equilibrium are called *canonical ensembles*. A simple example for the canonical ensemble is just a bottle of water put on the table. The energy of the system in canonical ensemble are changeable. Our ultimate goal in this subsection is to find a probability distribution of energy as for canonical ensembles.

First of all, systems in a canonical ensemble have the same as well as consistent temperature T as their reservoir. To explain this, we need to figure out what is T first.

Temperature

It turns out that entropy S is the core concept, and we normally view S as a function of (E, V, N) . Then the definition of temperature is

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{V, N} \quad (1.8)$$

To see why this is a correct definition, firstly we can deduce it from the basic formula of thermodynamics which is

$$dE = TdS - pdV + \mu dN$$

¹⁵Here we call the environment as *reservoir*, which is so large that all its macroscopic properties keep unchanged during the interacting with other little systems.

Then we simply modify this formula as

$$dS = \frac{1}{T}dE + \frac{p}{T}dV - \frac{\mu}{T}dN \quad (1.9)$$

From (1.9), we can directly get the (1.8). What's more, we can also get the (1.13) and (1.17) from it. (1.9) is just the *1st law of thermodynamics* which is rather familiar for us.

More importantly, there is another way to understand (1.8). We can consider a scenario below. Imagine an isolated system is divided by two parts —system *A* and

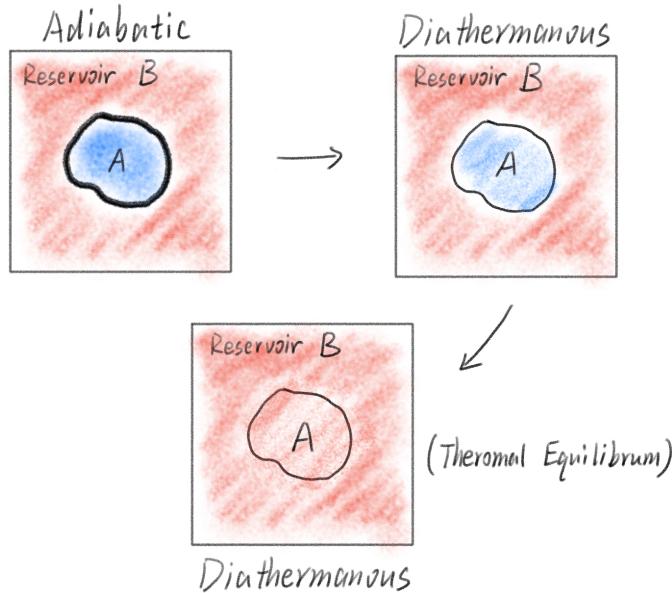


Figure 1.3: System A and Reservoir B

reservoir *B*. *A* is put inside of *B*. Initially, the cover of *A* is adiabatic, then we change the constraint, turning the cover of *A* into a diathermanous one. The equilibrium of the combined system is broken, and the whole system will evolve with time. Finally, it will reach a new equilibrium.

Suppose that *E* is the energy of *A*, while *E_{total}* is the total energy of *A* and *B*. *E_{total}* is a constant. Thus, the energy of *B* is *E_{total} - E*. It's not hard to figure out that the number of achievable microstates for the combined system is

$$\Omega_{\text{total}} = \sum \Omega_A \Omega_B = \sum_i \Omega_A(E_i) \Omega_B(E_{\text{total}} - E_i) = \sum_i e^{\frac{S_A(E_i) + S_B(E_{\text{total}} - E_i)}{k}}$$

from where we can know that Ω_{total} has a form of summation by lots of terms, and each one of these terms is corresponding to a macrostate for the combined system.

However, the magnitude of Ω_{total} is dominated by terms with largest magnitudes, and the real macrostates shown in the new equilibrium are the states corresponding to these terms. It's easy to know that the terms with largest magnitudes all satisfy that

$$\frac{\partial(\Omega_A \Omega_B)}{\partial E} \sim 0$$

i.e. we have

$$\frac{\partial}{\partial E} e^{\frac{S_A(E) + S_B(E_{\text{total}} - E)}{k}} = \frac{1}{k} \left(\frac{\partial S_A}{\partial E} - \frac{\partial S_B}{\partial E} \right) e^{\frac{S_A(E) + S_B(E_{\text{total}} - E)}{k}} \sim 0$$

thus we can know that $\frac{\partial S_A}{\partial E} \sim \frac{\partial S_B}{\partial E}$. Because there are only two parts in the combined system, so $|dE_A| = |dE_B|$. Hence we can get that the macrostate shown in thermal equilibrium of the system satisfy that

$$\frac{\partial S_A}{\partial E_A} \sim \frac{\partial S_B}{\partial E_B}$$

According to the definition (1.8), this means that during the new equilibrium we have

$$\frac{1}{T_A} \sim \frac{1}{T_B} \tag{1.10}$$

which is just the key characteristic of temperature. So that's why the definition of temperature (1.8) is reasonable¹⁶.

By the way, in thermodynamics (1.8) is usually called the condition of *thermal equilibrium*. Besides, we also have mechanical equilibrium and chemical equilibrium, the conditions of which will be seen below.

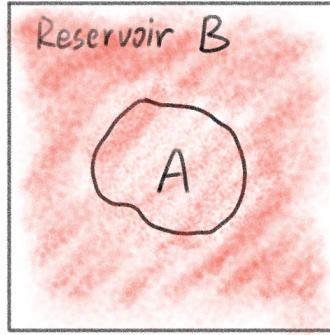
Probability Distribution

Now we can deduce the probability distribution of energy for a canonical ensemble. We denote the n -th microstate for A as $|E_n\rangle$, and its energy is E_n . The combined system A and B is in a microcanonical ensemble. Thus, we can calcu-

¹⁶The relation (1.10) is valid for any two systems, whether or not one of them is a reservoir. We can also have the intuition that although both of the systems' temperatures are changing with time, their ensemble averages are same, i.e. we indeed have $\langle T_A \rangle = \langle T_B \rangle$.

What's more, practically the difference of T_A and T_B is too small to be considered, thus we can directly write that $T_A = T_B = T$ without too many errors. Nonetheless, we still want to emphasize the uncertainty and the fluctuations, since those are the key characteristics of these large-consitituents systems and the main reason why we need the methods of statistics.

late the number of microstates for the combined system in equilibrium by using the formula of Boltzmann entropy as



$$\Omega_{\text{total}} = \sum_n \Omega_B (E_{\text{total}} - E_n) = \sum_n e^{\frac{S_B(E_{\text{total}} - E_n)}{k}}$$

Since B is a reservoir, E_{total} is much larger than ANY E_n , so we can do Taylor expansion to each terms above, hence the formula of Ω_{total} becomes

$$\Omega_{\text{total}} \approx \sum_n e^{\frac{S_B(E_{\text{total}})}{k} - \frac{1}{k} \frac{\partial S_B}{\partial E_B} E_n} = e^{\frac{S_B(E_{\text{total}})}{k}} \sum_n e^{-\frac{E_n}{kT}}$$

Then let's calculate the number of microstates that A stays in the microstate of $|E_n\rangle$, which is

$$\begin{aligned} \Omega_n &= \Omega_B (E_{\text{total}} - E_n) = e^{\frac{S_B(E_{\text{total}} - E_n)}{k}} \\ &\approx e^{\frac{S_B(E_{\text{total}})}{k} - \frac{1}{k} \frac{\partial S_B}{\partial E_B} E_n} \\ &= e^{\frac{S_B(E_{\text{total}})}{k}} \cdot e^{-\frac{E_n}{kT}} \end{aligned}$$

Since the combined system A and B is in a microcanonical ensemble, the probability to find the system A in microstate $|E_n\rangle$ is

$$p_n = \frac{\Omega_n}{\Omega_{\text{total}}} = \frac{e^{-\frac{E_n}{kT}}}{\sum_n e^{-\frac{E_n}{kT}}} \quad (1.11)$$

(1.11) is the probability distribution to find system in a canonical ensemble to stay in microstate $|E_n\rangle$. It's also called the *Boltzmann distribution*.

Subsequently, we can deduce the probability for the system to stay in the

macrostate of E as

$$p_E = \frac{\Omega(E)e^{-\frac{E}{kT}}}{\sum_n e^{-\frac{E_n}{kT}}} \quad (1.12)$$

Finally, I'd like to emphasize several things.

- Firstly, the summation $\sum_n e^{-\frac{E_n}{kT}}$ in the denominator above is over all the energy instead of over all the energy level. That's to say it can be a kind of case that $E_i = E_j (i \neq j)$.
- Secondly, as mentioned in one of the footnotes above, the T in the probability distribution is the temperature of the reservoir instead of the temperature of the system A, because the temperature of the A is changing with time. However, we can say that $\langle T_A \rangle = T$, and the fluctuation is rather small compared to the average itself.
- Finally, you should distinguish the microcanonical ensemble and the canonical ensemble here. The combined system A and B is in microcanonical ensemble, while only A itself is in canonical ensemble. Nonetheless, while deducing, we sometimes treat both A and B microcanonical ensembles, since we need to use their Ω and their Boltzmann entropies.

1.4.4 Grand Canonical Ensembles

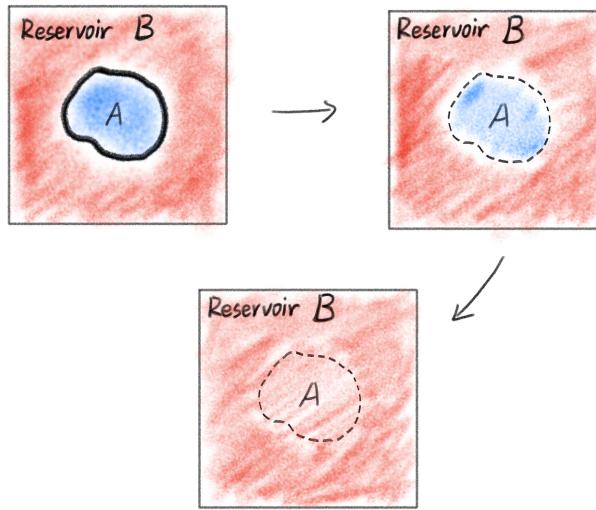
Then let's see another kind of ensembles, this time with fixed V , but the systems can have thermal interaction as well as constituents exchange with the reservoir. We call this kind of ensembles *grand canonical ensembles*. We want to find the probability distribution of N and E for grand canonical ensemble in this subsection.

We can set up a similar scenario, merely with a difference that after changing the constraint, the part A can have both thermal interaction and constituents exchange with part B . When the combined system reach new equilibrium, A and B will have same temperature and chemical potential.

Chemical Potential

The definition of chemical potential is

$$\mu = -T \left(\frac{\partial S}{\partial N} \right)_{E,V} \quad (1.13)$$



If we view S as a function of E and N , then similar analysis tells us that the new equilibrium for the combined system satisfy that

$$\begin{cases} \frac{\partial S_A}{\partial E_A} \sim \frac{\partial S_B}{\partial E_B} \\ \frac{\partial S_A}{\partial N_A} \sim \frac{\partial S_B}{\partial N_B} \end{cases}$$

which means $T_A \sim T_B$, which is the condition of thermal equilibrium, and the formula

$$\mu_A \sim \mu_B \quad (1.14)$$

which is the condition of *chemical equilibrium*.

Probability Distribution

Again, we can display a similar analysis and get that the probability to find part A in the microstate of $|E_{n,m}, N_m\rangle$ in new equilibrium is that

$$p_{n,m} = \frac{e^{-\frac{E_{n,m}}{kT} + \frac{\mu N_m}{kT}}}{\sum_{m,n} e^{-\frac{E_{n,m}}{kT} + \frac{\mu N_m}{kT}}} \quad (1.15)$$

and the probability distribution of E and N is

$$p_{E,N} = \frac{\Omega(E, N) e^{-\frac{E}{kT} + \frac{\mu N}{kT}}}{\sum_{m,n} e^{-\frac{E_{n,m}}{kT} + \frac{\mu N_m}{kT}}} \quad (1.16)$$

where the T and μ are the temperature and chemical potential of the reservoir.

1.4.5 Gibbs Ensembles and Pressure

Finally, let's consider another kind of ensemble, whose systems have fixed N , and can have thermal as well as mechanical interaction with the environment. We name this kind of ensemble as Gibbs ensemble.

It's not hard to figure out that during equilibrium, the system will have the same magnitude of pressure as the environment.

Pressure

The definition of pressure is

$$p = T \left(\frac{\partial S}{\partial V} \right)_{E,N} \quad (1.17)$$

By using the same method, we can analyze that the mechanical equilibrium condition is

$$p_A \sim p_B \quad (1.18)$$

As for the probability distribution for Gibbs ensemble, since I am a little bit tired, and lazy, so I want to leave this task for you.

1.4.6 How to Understand the Names of Different Ensembles?

Up to now, many of you might be confused about how to understand the name of these ensembles literally? Especially for the microcanonical, canonical and grand canonical ensembles, what's their names meaning?

Well, I have thought a way to understand.

- First of all, both the canonical ensemble and the grand canonical ensemble can be viewed as under the probability distribution of different microcanonical ensembles with different E and N . In this sense, the microcanonical ensembles play the same role as microstates, aren't they? So that's why there is a "micro" in its name I suppose.
- Then, "canonical" means the most common and the most natural in some extent. Since the canonical ensemble describes the most common and the most natural situation, there is "canonical" in its name.

- As for the grand canonical ensemble, it has "grand" in the name because there are much more kinds of microcanonical ensembles in the probability distribution.

The arguments above can provide a good way to generalize the concept of entropy to other kinds of systems(ensembles). We will see that next section.

1.5 Partiton Function

1.5.1 Partition Function

Partition function is the core tool in statistical mechanics. The definition of partition function for canonical ensemble is

$$Z = \sum_n e^{-\frac{E_n}{kT}} = \sum_n e^{-\beta E_n} \quad (1.19)$$

which is just the denominator of (1.11). Here, we also denote that

$$\beta = \frac{1}{kT} \quad (1.20)$$

By using this notation, we can write the probability distribution for canonical ensemble as

$$p_n = \frac{e^{-\beta E_n}}{Z} \quad (1.21)$$

Similarly, for grand canonical ensemble, we have the *grand canonical partition function*

$$\Xi = \sum_{m,n} e^{-\beta(E_{n,m} - \mu N_m)} \quad (1.22)$$

Why Z is named as "partition function"? According to (1.21), we can know that Z determines how energy is partitioned to the system —this formula reveals the probability distribution of the way to partition energy.

Finally, let's have a look at a useful property for partition function. Suppose that there are two independent systems labeled by A and B with same temperature. Both of them are canonical ensembles, so we can calculate the partition function of the combined system A and B as

$$Z_C = \sum_{a,b} e^{-\beta(E_a + E_b)} = \sum_a e^{-\beta E_a} \cdot \sum_b e^{-\beta E_b} = Z_A Z_B \quad (1.23)$$

That's to say, the partition function of the combined system equals to the multiplication of the partition functions of its independent constituents.

1.5.2 Simple Applications of Partition Function

In this subsection, I want to show that the partition function can be our best friend when addressing the problem of systems that we care interested in. That's

because, once we have the concrete form of the partition function of a system, we can deduce almost all the other useful quantities by using it. Let's see how to achieve that.

1. Averages of E and N

As is mentioned for a lot of times, the system in equilibrium is jumping among different microstates. Hence all the measurements will have results with a sense of average.

If we want to calculate the average energy of a system in a canonical ensemble, we can use the formula

$$\langle E \rangle = \sum_n p_n E_n = \frac{\sum_n E_n e^{-\beta E_n}}{\sum_n e^{-\beta E_n}}$$

If you stare at this formula carefully for a while, you will find that

$$\langle E \rangle = -\frac{\frac{\partial}{\partial \beta} (\sum_n e^{-\beta E_n})}{\sum_n e^{-\beta E_n}} = -\frac{\frac{\partial}{\partial \beta} Z}{Z} = -\frac{\partial}{\partial \beta} \ln Z \quad (1.24)$$

That's exactly what we want.

Similarly, for grand canonical ensemble, it's easy to know that

$$-\frac{\partial}{\partial \beta} \ln \Xi = \langle E \rangle - \mu \langle N \rangle$$

Considering that it's easy to get the average of N as

$$\langle N \rangle = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \Xi \quad (1.25)$$

then we can deduce that

$$\langle E \rangle = -\frac{\partial}{\partial \beta} \ln \Xi + \frac{\mu}{\beta} \frac{\partial}{\partial \mu} \ln \Xi \quad (1.26)$$

2. Dispersions(Variances) of E and N

After a brief discussion on average, next we can have a look at the dispersion, from which we can conclude something interesting for systems with a large number of constituents.

First of all, the dispersion of energy for system in canonical ensemble can be

calculated as

$$\begin{aligned}
 \sigma_E^2 &= \langle E^2 \rangle - \langle E \rangle^2 \\
 &= \sum_n E_n^2 \frac{e^{-\beta E_n}}{Z} - \left(\sum_n E_n \frac{e^{-\beta E_n}}{Z} \right)^2 \\
 &= \frac{\frac{\partial^2}{\partial \beta^2} Z}{Z} - \left(\frac{\frac{\partial}{\partial \beta} Z}{Z} \right)^2 \\
 &= \frac{Z \frac{\partial^2}{\partial \beta^2} Z - \frac{\partial Z}{\partial \beta} \frac{\partial Z}{\partial \beta}}{Z^2} \\
 &= \frac{\partial^2}{\partial \beta^2} \ln Z \\
 &= -\frac{\partial \langle E \rangle}{\partial \beta}
 \end{aligned} \tag{1.27}$$

Then, for grand canonical ensembles, it might be a little complex to calculate σ_E^2 . Firstly, let consider that

$$\begin{aligned}
 \frac{\partial}{\partial \beta} \langle E \rangle &= \frac{\sum_{m,n} -(E_n - \mu N_m) E_n e^{-\beta(E_n, m - \mu N_m)}}{\Xi} \\
 &\quad - \frac{\sum_{m,n} E_n e^{-\beta(E_n, m - \mu N_m)}}{\Xi} \cdot \frac{\sum_{m,n} -(E_n - \mu N_m) e^{-\beta(E_n, m - \mu N_m)}}{\Xi} \\
 &= -\langle E^2 - \mu EN \rangle + \langle E \rangle (\langle E \rangle - \mu \langle N \rangle) \\
 &= -\langle E^2 \rangle + \mu \langle EN \rangle + \langle E \rangle^2 - \mu \langle E \rangle \langle N \rangle \\
 &= \mu \text{Cov}(E, N) - \sigma_E^2
 \end{aligned}$$

where the $\text{Cov}(E, N) = \langle EN \rangle - \langle E \rangle \langle N \rangle$ is the *covariance* we learnt in statistics. Then, let's calculate that

$$\begin{aligned}
 \frac{\partial}{\partial \beta} \langle N \rangle &= \frac{\partial}{\partial \beta} \frac{\sum_{m,n} N_m e^{-\beta(E_n, m - \mu N_m)}}{\Xi} \\
 &= \frac{\sum_{m,n} -N_m (E_n - \mu N_m) e^{-\beta(E_n, m - \mu N_m)}}{\Xi} \\
 &\quad - \frac{\sum_{m,n} N_m e^{-\beta(E_n, m - \mu N_m)}}{\Xi} \cdot \frac{\sum_{m,n} -(E_n - \mu N_m) e^{-\beta(E_n, m - \mu N_m)}}{\Xi} \\
 &= -\langle NE - \mu N^2 \rangle + \langle N \rangle \langle E - \mu N \rangle \\
 &= -\langle NE \rangle + \langle N \rangle \langle E \rangle + \mu \langle N^2 \rangle - \mu \langle N \rangle^2 \\
 &= \mu \sigma_N^2 - \text{Cov}(N, E)
 \end{aligned}$$

Thus we have

$$\mu\sigma_N^2 - \frac{\partial\langle N \rangle}{\partial\beta} = \frac{1}{\mu} \left(\sigma_E^2 + \frac{\partial\langle E \rangle}{\partial\beta} \right)$$

i.e.

$$\sigma_E^2 = -\frac{\partial\langle E \rangle}{\partial\beta} - \mu \frac{\partial\langle N \rangle}{\partial\beta} + \mu^2 \sigma_N^2$$

It's easy to know that the dispersion of N is

$$\begin{aligned} \sigma_N^2 &= \langle N^2 \rangle - \langle N \rangle^2 \\ &= \frac{\sum_{m,n} N_m^2 e^{-\beta(E_{n,m} - \mu N_m)}}{\Xi} - \left[\frac{\sum_{m,n} N_m e^{-\beta(E_{n,m} - \mu N_m)}}{\Xi} \right]^2 \\ &= \frac{\frac{1}{\beta^2} \frac{\partial^2}{\partial\mu^2} \Xi}{\Xi} - \frac{\frac{1}{\beta^2} \left(\frac{\partial\Xi}{\partial\mu} \right)^2}{\Xi^2} \\ &= \frac{1}{\beta^2} \frac{\partial^2}{\partial\mu^2} \ln \Xi \end{aligned} \tag{1.28}$$

Thus we ultimately have the formula of σ_E^2 for grand canonical ensemble as

$$\begin{aligned} \sigma_E^2 &= -\frac{\partial\langle E \rangle}{\partial\beta} - \mu \frac{\partial\langle N \rangle}{\partial\beta} + \mu^2 \sigma_N^2 \\ &= -\frac{\partial}{\partial\beta} \left(-\frac{\partial}{\partial\beta} \ln \Xi + \frac{\mu}{\beta} \frac{\partial}{\partial\mu} \ln \Xi \right) - \mu \frac{\partial}{\partial\beta} \left(\frac{1}{\beta} \frac{\partial}{\partial\mu} \ln \Xi \right) + \mu^2 \cdot \frac{1}{\beta^2} \frac{\partial^2}{\partial\mu^2} \ln \Xi \\ &= \frac{\partial^2}{\partial\beta^2} \ln \Xi + \frac{2\mu}{\beta^2} \frac{\partial}{\partial\mu} \ln \Xi - \frac{2\mu}{\beta} \frac{\partial^2}{\partial\mu\partial\beta} \ln \Xi + \frac{\mu^2}{\beta^2} \frac{\partial^2}{\partial\mu^2} \ln \Xi \end{aligned} \tag{1.29}$$

Thermodynamic Limit

Now let's consider a system in canonical ensemble with N independent¹⁷ particles in it. Suppose the partition function of each particle is Z , thus according to the property of partition function, the partition function of the whole system is Z^N ¹⁸.

Then let's calculate the fraction of σ_E to $\langle E \rangle$ and the fraction of σ_N to $\langle N \rangle$.

¹⁷ Here independent can simply be understood as there are no interactions between these particles, i.e. the system is ideal. However, for the real systems exist in nature, of course the particles aren't independent. But the conclusions we get below can also be valid in some extent, since what we get is the main property for the system.

Phycis is that. Sometimes you need to grab the main characteristic and properly ignore the other secondary cause, nonetheless the results you get won't be so far away from the truth.

¹⁸ Actually, this form of the whole partition function is not correct. We get this form since we implicitly assume that all the particles in the system are distinguishable. But for the system consists of the same kind of particles, the particles are essentially and naturally indistinguishable, thus leading to a modification of the formula of the partition function. We will discuss this in next chapter where we argue the Gibbs paradox. For now, you don't need to be so worry because although we take the modification into account, the results and conclusions are the same.

Firstly, for canonical ensemble, the fraction is

$$\frac{\sigma_E}{\langle E \rangle} = \frac{\sqrt{\frac{\partial^2}{\partial \beta^2} \ln Z^N}}{-\frac{\partial}{\partial \beta} \ln Z^N} = \frac{1}{\sqrt{N}} \frac{\sqrt{\frac{\partial^2}{\partial \beta^2} \ln Z}}{-\frac{\partial}{\partial \beta} \ln Z} \sim \frac{1}{\sqrt{N}}$$

So we have that $\lim_{N \rightarrow \infty} \frac{\sigma_E}{\langle E \rangle} = 0$.

Furthermore, let's consider a system in grand canonical ensemble. This time we can't say that the system has N particles in it because the number of particles is changable. However, we can also get the similar conclusion for the grand canonical ensemble that $\lim_{\langle N \rangle \rightarrow \infty} \frac{\sigma_E}{\langle E \rangle} = \lim_{\langle N \rangle \rightarrow \infty} \frac{\sigma_N}{\langle N \rangle} = 0$. To illustrate that, we only need to consider the grand canonical partition function Ξ and modify it as

$$\Xi = \sum_{n,m} e^{-\beta(E_{n,m} - \mu N_m)} = \sum_m e^{\beta \mu N_m} \sum_n e^{-\beta E_{n,m}} = \sum_m e^{\beta \mu N_m} Z^{N_m}$$

where Z is still the partion function for a single particle. With this form of formula, now its reasonable why we said that $\lim_{\langle N \rangle \rightarrow \infty} \frac{\sigma_E}{\langle E \rangle} = \lim_{\langle N \rangle \rightarrow \infty} \frac{\sigma_N}{\langle N \rangle} = 0$ is still valid because we have N_m as the power.

The condition $\langle N \rangle \rightarrow \infty$ is called the *thermodynamic limit*. According to the conclusions above we can know that under the thermodynamic limit, the fluctuation of E and N of a system interacting with environment are ignorable compared to the average of E and N ¹⁹. Thus, under the thermodynamic limit, microcanonical, canonical and grand canonical ensembles have same properties in this sense. Therefore, we can replace $\langle E \rangle$ and $\langle N \rangle$ by E and N from now on if the system has a very large number of constituents.

Besides the E and N , we can further argue that *all the macroscopic properties are same for different ensembles under thermodynamic limit*. This is an important property for the system that we are studying, and we can easily understand it by the way below.

Now we label the i -th energy level with E_i , and there might be lots of microstates

¹⁹This conclusion, in my point of view, has a direrctly relation with the *central limit theorem* in statistics.

have the same energy of E_i ²⁰. Then the partition function for canonical ensemble is

$$Z = \sum_i \Omega(E_i) e^{-\beta E_i}$$

Since we already know that under the thermodynamic limit $\frac{\Delta E}{\langle E \rangle} \sim 0$, thus the number of microstates with energy of $\langle E \rangle$, i.e. $\Omega(\langle E \rangle)$, is dominative. Then we can reasonably write that

$$Z = \Omega(\langle E \rangle) e^{-\beta \langle E \rangle}$$

Similarly, for grand canonical ensembles, under thermodynamic limit, we have

$$\Xi = \Omega(\langle E \rangle, \langle N \rangle) e^{-\beta(\langle E \rangle - \mu \langle N \rangle)}$$

These forms of partition functions are exactly the same as the form of partition function of microcanonical ensembles.

Now that, as we will immediately see below, all the things can be deduced by the partition functions, and under thermodynamic limit, all kinds of partition functions coincide with ignorable errors. Thus in equilibrium, all the conclusions, including all the forms of formulas, deduced from one kind of ensemble can automatically be valid for the other kinds of ensembles as long as we have the thermodynamic limit.

3. Average of Pressure

Actually, the pressure of a system in both canonical and grand canonical ensembles are fluctuating with time. Thus the pressure it shows oughts to be the average.

For canonical ensemble, the average pressure is

$$\langle p \rangle = \sum_n -\frac{\partial E_n}{\partial V} \frac{e^{-\beta E_n}}{Z} = \frac{1}{Z} \sum_n \frac{1}{\beta} \frac{\partial}{\partial V} e^{-\beta E_n} = \frac{1}{\beta} \frac{\partial \ln Z}{\partial V} \quad (1.30)$$

while for grand canonical ensemble, the average pressure is

$$\langle p \rangle = \sum_{m,n} -\frac{\partial E_n}{\partial V} \frac{e^{-\beta(E_{n,m} - \mu N_m)}}{\Xi} = \frac{1}{\Xi} \sum_{n,m} \frac{1}{\beta} \frac{\partial}{\partial V} e^{-\beta(E_{n,m} - \mu N_m)} = \frac{1}{\beta} \frac{\partial \ln \Xi}{\partial V} \quad (1.31)$$

4. Entropy

We have met the definition of entropy for an isolated system, or in other words, for the micro-canonical ensemble. We have also known that all kinds of systems can

²⁰This is different from what we used to do. Before we labeled the energy of microstates instead of the energy levels, thus at that time we didn't consider the degeneracy.

be viewed as sitting in the probability distribution of isolated systems with different states. Thus the entropy of a system also has the sense of average. Now let's deduce the average entropy by using the average version of basic relation of thermodynamics.

For canonical ensembles, we have

$$\begin{aligned}
 d\langle S \rangle &= \frac{1}{T} d\langle E \rangle + \frac{\langle p \rangle}{T} dV \\
 &= k\beta d\left(-\frac{\partial}{\partial\beta} \ln Z\right) - k\frac{\partial}{\partial V} \ln Z dV \\
 &= k d\left(-\beta\frac{\partial}{\partial\beta} \ln Z\right) + k\frac{\partial}{\partial\beta} \ln Z d\beta + k\frac{\partial}{\partial V} \ln Z dV \\
 &= d\left(-k\beta\frac{\partial}{\partial\beta} \ln Z + k \ln Z\right) \\
 &= d\left(\frac{1}{T} \cdot kT^2 \frac{\partial}{\partial T} \ln Z + k \ln Z\right) \\
 &= d\left[\frac{\partial}{\partial T}(kT \ln Z)\right]
 \end{aligned}$$

where T is the temperature of the reservoir. Thus the average entropy of canonical ensembles is

$$\langle S \rangle = -k\beta \frac{\partial}{\partial\beta} \ln Z + k \ln Z = \frac{\partial}{\partial T}(kT \ln Z) \quad (1.32)$$

Similarly, for grand canonical ensembles, we have

$$\begin{aligned}
 d\langle S \rangle &= \frac{1}{T} d\langle E \rangle + \frac{\langle p \rangle}{T} dV - \frac{\mu}{T} d\langle N \rangle \\
 &= k\beta d\left(-\frac{\partial}{\partial\beta} \ln \Xi + \frac{\mu}{\beta} \frac{\partial}{\partial\mu} \ln \Xi\right) + k\frac{\partial}{\partial V} \ln \Xi dV - k\beta\mu d\left(\frac{1}{\beta} \frac{\partial}{\partial\mu} \ln \Xi\right) \\
 &= d\left(-k\beta\frac{\partial}{\partial\beta} \ln \Xi + k\mu\frac{\partial}{\partial\mu} \ln \Xi\right) + k\frac{\partial}{\partial\beta} \ln \Xi d\beta \\
 &\quad - k\frac{\mu}{\beta} \frac{\partial}{\partial\mu} \ln \Xi d\beta + k\frac{\partial}{\partial V} \ln \Xi dV - d\left(k\mu\frac{\partial}{\partial\mu} \ln \Xi\right) + k\frac{1}{\beta} \frac{\partial}{\partial\mu} \ln \Xi d(\beta\mu) \\
 &= d\left(-k\beta\frac{\partial}{\partial\beta} \ln \Xi\right) + k\frac{\partial}{\partial\beta} \ln \Xi d\beta + k\frac{\partial}{\partial V} \ln \Xi dV + k\frac{\partial}{\partial\mu} \ln \Xi d\mu \\
 &= d\left(-k\beta\frac{\partial}{\partial\beta} \ln \Xi + k \ln \Xi\right) \\
 &= d\left[\frac{\partial}{\partial T}(kT \ln \Xi)\right]
 \end{aligned}$$

where T and μ are temperature and chemical potential of the reservoir. Thus the

average entropy of grand canonical ensembles is

$$\langle S \rangle = -k\beta \frac{\partial}{\partial \beta} \ln \Xi + k \ln \Xi = \frac{\partial}{\partial T} (kT \ln \Xi) \quad (1.33)$$

5. Free Energy

The definition of *free energy* is

$$F = E - TS \quad (1.34)$$

Still, if we are talking about the canonical ensemble or the grand canonical ensemble, the T is the temperature of the reservoir. Naturally, F oughts to be viewed as a function of (T, V, N) , since we have

$$\begin{aligned} dF &= dE - d(TS) \\ &= TdS - pdV + \mu dN - TdS - SdT \\ &= -SdT - pdV + \mu dN \end{aligned}$$

Hence we can calculate S, p and μ from F easily as

$$\begin{cases} S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \\ p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \\ \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \end{cases} \quad (1.35)$$

The reason why we particularly pick out F is that F has the most direct relationship with the partition function.

For canonical ensembles, the free energy also should be viewed as in a sense of average, and we can deduce that

$$\begin{aligned} \langle F \rangle &= \langle E \rangle - T\langle S \rangle \\ &= -\frac{\partial}{\partial \beta} \ln Z - T \left(-k\beta \frac{\partial}{\partial \beta} \ln Z + k \ln Z \right) \\ &= -kT \ln Z \end{aligned} \quad (1.36)$$

And for grand canonical ensembles, we can deduce that

$$\begin{aligned} \langle F \rangle &= \langle E \rangle - T\langle S \rangle \\ &= -\frac{\partial}{\partial \beta} \ln \Xi + \frac{\mu}{\beta} \frac{\partial}{\partial \mu} \ln \Xi - T \left(-k\beta \frac{\partial}{\partial \beta} \ln \Xi + k \ln \Xi \right) \end{aligned}$$

$$= -\frac{1}{\beta} \ln \Xi + \frac{\mu}{\beta} \frac{\partial}{\partial \mu} \ln \Xi \quad (1.37)$$

Last but not least, although complex, it can be proved by similar methods that under the thermodynamic limit, all the average above is practically much larger than their fluctuations. So we can happily throw out all the average symbols from now on if we are dealing with systems with large number of constituents.

Chapter 2

Ideal Classical Gases

In this chapter, what we are interested in is the ideal classical gases consisting of N identical molecules. Here ideal means the molecules are independent and have no interactions with each other. Our strategy is that we first view the system populates in the canonical ensemble. *Since the system is under the thermodynamic limit, the results we get can be finely valid for systems in microcanonical or grand canonical ensembles.*

2.1 Classical Partition Function

2.1.1 Classical Partition Function

Actually, the happiest thing in statistical mechanics is finding the partition functions. Now let's deduce the form of classical partition function for a single particle from its quantum form, where we view the classical form as the classical limit form of the quantum form.

For the canonical ensemble, it's easy to know that the partition function for a single particle is

$$Z = \sum_n e^{-\beta E_n} = \sum_n \langle n | e^{-\beta \hat{H}} | n \rangle = \text{Tr } e^{-\beta \hat{H}}$$

where $|n\rangle$ is the orthonormal energy eigenstate of the particle, and \hat{H} is the Hamiltonian operator that contains terms of all the degrees of freedom. Above we have chosen the energy representation. However, the trace of an operator is independent of the choice of representation¹. That's to say, we can also represent Z by using the position representation as

$$Z = \int_{-\infty}^{+\infty} \langle \mathbf{q} | e^{-\beta \hat{H}} | \mathbf{q} \rangle d^n \mathbf{q}$$

where $|\mathbf{q}\rangle$ is the position eigenstate and n is the degree of freedom. You need to keep in mind that here \mathbf{q} contains all the general coordinates instead of only for the spatial coordinates.

Then, let's consider a useful relation between operators.

$$\begin{aligned} e^{\hat{A}} e^{\hat{B}} &= \left(1 + \hat{A} + \frac{\hat{A}^2}{2!} + \dots \right) \left(1 + \hat{B} + \frac{\hat{B}^2}{2!} + \dots \right) \\ &= 1 + (\hat{A} + \hat{B}) + \frac{1}{2!}(\hat{A} + \hat{B})^2 + \dots + \frac{1}{n!}(\hat{A} + \hat{B})^n + o([\hat{A}, \hat{B}]) \end{aligned}$$

¹This is an important theorem in quantum mechanics.

$$= e^{\hat{A}+\hat{B}} + o[\hat{A}, \hat{B}]$$

So in our context now, we can get

$$e^{-\beta\hat{T}(\hat{\mathbf{p}})}e^{-\beta\hat{V}(\hat{\mathbf{q}})} = e^{-\beta\hat{H}} + \beta o([\hat{q}, \hat{p}]) = e^{-\beta\hat{H}} + \beta o(\hbar)$$

where for all the general coordinates and their conjugate momentum, the commutative relation $[\hat{q}, \hat{p}] = i\hbar$ is valid thanks to the *canonical quantizations*. Then because that under the classical limit, we have $\hbar \sim 0$, thus we have²

$$e^{-\beta\hat{H}} = e^{-\beta\hat{T}(\hat{\mathbf{p}})}e^{-\beta\hat{V}(\hat{\mathbf{q}})} \quad (2.1)$$

Hence, we can calculate the partition for a single particle under the classical limit as

$$\begin{aligned} Z &= \int_{-\infty}^{+\infty} \langle \mathbf{q} | e^{-\beta\hat{H}} | \mathbf{q} \rangle d^n \mathbf{q} \\ &= \int_{-\infty}^{+\infty} \langle \mathbf{q} | e^{-\beta\hat{T}(\hat{\mathbf{p}})} e^{-\beta\hat{V}(\hat{\mathbf{q}})} | \mathbf{q} \rangle d^n \mathbf{q} \\ &= \int_{-\infty}^{+\infty} e^{-\beta V(\mathbf{q})} \langle \mathbf{q} | e^{-\beta\hat{T}(\hat{\mathbf{p}})} | \mathbf{q} \rangle d^n \mathbf{q} \\ &= \iiint_{-\infty}^{+\infty} e^{-\beta V(\mathbf{q})} \langle \mathbf{q} | \mathbf{p}' \rangle \langle \mathbf{p}' | e^{-\beta\hat{T}(\hat{\mathbf{p}})} | \mathbf{p} \rangle \langle \mathbf{p} | \mathbf{q} \rangle d^n \mathbf{q} d^n \mathbf{p}' d^n \mathbf{p} \\ &= \iiint_{-\infty}^{+\infty} e^{-\beta H(\mathbf{q}, \mathbf{p})} \langle \mathbf{q} | \mathbf{p}' \rangle \langle \mathbf{p}' | \mathbf{p} \rangle \langle \mathbf{p} | \mathbf{q} \rangle d^n \mathbf{q} d^n \mathbf{p}' d^n \mathbf{p} \\ &= \iiint_{-\infty}^{+\infty} e^{-\beta H(\mathbf{q}, \mathbf{p})} \langle \mathbf{p}' | \mathbf{p} \rangle \langle \mathbf{p} | \mathbf{q} \rangle \langle \mathbf{q} | \mathbf{p}' \rangle d^n \mathbf{q} d^n \mathbf{p}' d^n \mathbf{p} \\ &= \iint_{-\infty}^{+\infty} e^{-\beta H(\mathbf{q}, \mathbf{p})} \langle \mathbf{p} | \mathbf{q} \rangle \langle \mathbf{q} | \mathbf{p} \rangle d^n \mathbf{q} d^n \mathbf{p} \end{aligned}$$

Since according to quantum mechanics we have $\langle \mathbf{p} | \mathbf{q} \rangle = \left(\frac{1}{2\pi\hbar}\right)^{\frac{n}{2}} e^{\frac{i\mathbf{p} \cdot \mathbf{q}}{\hbar}} = \langle \mathbf{q} | \mathbf{p} \rangle^*$, so we have $\langle \mathbf{p} | \mathbf{q} \rangle \langle \mathbf{q} | \mathbf{p} \rangle = \frac{1}{h^n}$, where h is the *Plank constant*. Thus, ultimately, we get

$$Z = \iint_{-\infty}^{+\infty} e^{-\beta H(\mathbf{q}, \mathbf{p})} \frac{d^n \mathbf{q} d^n \mathbf{p}}{h^n} \quad (2.2)$$

²Or you can just simply think in this way: under the classical limit, $[\hat{q}, \hat{p}] = i\hbar \sim 0$. Thus \hat{q} and \hat{p} are simply commutative, which means we can deduce (2.1) directly.

A Classical Way of Thinking

Finally, there is a classical as well as intuitive way to understand (2.2). If you insist that you want to get the classical partition function without any quantum things, you can use this way of thinking.

We know that one of the main differences between classical world and quantum world is that the former is continuous while the latter is discrete. Thus, when transitioning the partition function from quantum form to classical form, we should replace the discrete summation by continuous integral.

Actually, this kind of integral is over the *phase cell*. We divide the whole phase space of the particle by little phase cells, whose "volume" equals to \hbar^n . And in the classical context, we chose one point in a phase cell to represent a microstate of the particle. By thinking as this way, we can understand (2.2) properly.

We will also need to know how to understand the Boltzmann distribution in the classical world. We will meet this in section 2.3.

2.1.2 Monatomic Molecules

Then, let's consider the concrete examples for our purposes of this chapter. Suppose we have a container, whose volume is V . This container has N ideal monatomic molecules in it. Ideal means that the Hamiltonian of a single monatomic molecule is

$$H = \frac{p^2}{2m} = \frac{p_x^2 + p_y^2 + p_z^2}{2m}$$

and we don't have the potential terms. What's more above we only consider the translation mode. Thus, we can calculate the partition function of a single particle

$$\begin{aligned} Z &= Z_x Z_y Z_z \\ &= \frac{1}{\hbar^3} \int_{-\infty}^{+\infty} d^3 \hat{q} \int_{-\infty}^{+\infty} e^{-\beta \frac{p^2}{2m}} d^3 \mathbf{p} \\ &= \frac{V}{\hbar^3} \iiint_{-\infty}^{+\infty} e^{-\beta \frac{p_x^2 + p_y^2 + p_z^2}{2m}} dp_x dp_y dp_z \end{aligned}$$

The x component of the integral above has the form that

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{-\frac{\beta}{2m}p_x^2} dp_x &= \sqrt{\frac{2m}{\beta}} \int_{-\infty}^{+\infty} e^{-\sqrt{\frac{\beta}{2m}}p_x^2} d\sqrt{\frac{\beta}{2m}}p_x \\ &= \sqrt{2mkT} \int_{-\infty}^{+\infty} e^{-t^2} dt \\ &= \sqrt{2\pi mkT} \end{aligned}$$

By the way, if you don't know how to deal with the last integral, you can have a look at the appendix A.

So finally, the partition function for a single monatomic molecule is

$$Z = (2\pi mkT)^{\frac{3}{2}} \frac{V}{h^3} \quad (2.3)$$

2.1.3 Polyatomic Molecules

Above , since we are talking about the gas consists of monatomic molecule, we only need to consider the 3 translation degrees of freedom in the Hamiltonian. However, if we are dealing with gas that consists of polyatomic molecules, we must take the degrees of freedom of rotation and vibration into account.

For polyatomic molecules, the translation partition function are the same as monatomic molecule's. So below we only manage to find the properties of the rotation and the vibration partition functions. *Since all the degrees of freedom are independent, so we can research their partition function one by one and finally multiply all the partition functions to get the partition function for the whole molecule.*

By the way, if you have trouble in deducing the Hamiltonians below, you'd better revisit the knowledge in theoretical mechanics and Hamiltonian dynamics.

Rotation

Generally speaking , the degree of freedom for rotation is 3, and we choose the *Euler angles* (θ, ϕ, ψ) as general coordinates to describe the rotation of a polyatomic molecule. Furthermore , by choosing 3 *main axes* , we can write the rotation component of the molecule as

$$H_{\text{rot}} = \frac{L_1^2}{2I_1} + \frac{L_2^2}{2I_2} + \frac{L_3^2}{2I_3}$$

where L_i and I_i are the angular momentum and moment of inertia around the i -th main axis. Subsequently, let's use the conjugate momentum p_θ, p_ϕ and p_ψ to represent

L_1, L_2 and L_3 as

$$\begin{cases} L_1 = \frac{\sin \psi}{\sin \theta} (p_\phi - p_\psi \cos \theta) + p_\theta \cos \psi \\ L_2 = \frac{\cos \psi}{\sin \theta} (p_\phi - p_\psi \cos \theta) - p_\theta \sin \psi \\ L_3 = p_\psi \end{cases}$$

Now we can calculate the partition function of rotation as

$$Z_{\text{rot}} = \frac{1}{h^3} \iint_{-\infty}^{+\infty} e^{-\beta H_{\text{rot}}} d^3 \mathbf{q} d^3 \mathbf{p}$$

To calculate this integral, we can change the volume element $d^3 \mathbf{p} = dp_\theta dp_\phi dp_\psi$ into $d^3 \mathbf{L} = dL_1 dL_2 dL_3$ by multiply a *Jacobi factor* $J = \frac{\partial(p_\theta, p_\phi, p_\psi)}{\partial(L_1, L_2, L_3)}$. Simple calculation gives us that here $J = \sin \theta$. Thus, the partition function becomes

$$\begin{aligned} Z_{\text{rot}} &= \frac{1}{h^3} \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi \int_0^{2\pi} d\psi \int_{-\infty}^{+\infty} e^{-\beta \left(\frac{L_1^2}{2I_1} + \frac{L_2^2}{2I_2} + \frac{L_3^2}{2I_3} \right)} dL_1 dL_2 dL_3 \\ &= \frac{8\pi^2}{h^3} \sqrt{I_1 I_2 I_3} (2\pi kT)^{\frac{3}{2}} \\ &\propto T^{\frac{3}{2}} \end{aligned} \tag{2.4}$$

However, if the molecule is linear, then its degree of freedom of rotation reduces to 2, and we only have θ and ϕ as the general coordinates. So we let $\psi = 0$, and can have the relation $L_1 = p_\theta$, $L_2 = \frac{p_\phi}{\sin \theta}$. What's more, we also reasonably have $I_1 = I_2 = I$ for a linear molecule. Thus, the rotation Hamiltonian ultimately becomes

$$H_{\text{rot}} = \frac{L_1^2 + L_2^2}{2I} = \frac{p_\theta^2}{2I} + \frac{p_\phi^2}{2I \sin^2 \theta}$$

Then we can calculate the rotation partition function as

$$\begin{aligned} Z_{\text{rot}} &= \frac{1}{h^2} \int d^2 \mathbf{q} \int_{-\infty}^{+\infty} e^{-\beta \left(\frac{p_\theta^2}{2I} + \frac{p_\phi^2}{2I \sin^2 \theta} \right)} dp_\theta dp_\phi \\ &= \frac{2\pi I kT}{h^2} \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi \\ &= \frac{8\pi^2 I kT}{h^2} \\ &\propto T^{\frac{3}{2}} \end{aligned} \tag{2.5}$$

As for monatomic molecules, it has no degree of freedom of rotation, thus doesn't have the corresponding partition function.

Vibration

We know that in a molecule, atoms are connected by chemical bonds, so the atoms can't fly away from the molecule, but can only vibrate around their equilibrium position. For a molecule consists of N atoms, the total degree of freedom is $3N$. So the degree of freedom of vibration for the whole molecule equals to $3N - t - r$, where t and r are degrees of freedom of translation and rotation respectively. Then we simply view every degree of freedom of vibration performing as a harmonic oscillator with the frequency of ω .

Each degree of freedom of vibration has the form of Hamiltonian as

$$H_{\text{vib}} = \frac{p_\xi^2}{2\mu} + \frac{1}{2}\mu\omega^2\xi^2$$

where μ is something like "general mass" and $p_\xi = \dot{\mu\xi}$ is the conjugate momentum of the general coordinate ξ . Thus, the vibration partition function of this vibration degree of freedom is

$$\begin{aligned} Z_{\text{vib}} &= \frac{1}{h} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\beta\left(\frac{p_\xi^2}{2\mu} + \frac{\mu\omega^2\xi^2}{2}\right)} d\xi dp_\xi \\ &= \frac{1}{h} \cdot \sqrt{2\pi\mu kT} \cdot \sqrt{\frac{2\pi kT}{\mu\omega^2}} \\ &= \frac{kT}{\hbar\omega} \\ &\propto T^{\frac{3}{2}} \end{aligned} \tag{2.6}$$

The result above is merely for a single vibration. If there are v degree of freedom of vibration in a molecule, then the whole vibration partition function Z_{vib} oughts to have the property that $Z_{\text{vib}} \propto T^{\frac{3v}{2}}$.

Conclusion

According to the arguments above, we can easily conclude that, generally, for one molecule in the ideal classical gas , its partition function has the property that

$$Z = Z_{\text{trans}}Z_{\text{rot}}Z_{\text{vib}} \propto VT^{\frac{t+r+2v}{2}} \tag{2.7}$$

where t, r, v are the degrees of freedom of translation , rotation and vibration respectively. Here we haven't found the exact form of Z , because the property above is sufficient for us to proceed.

Above we finely find the partition function for a single molecule. But, what about the partition function of the whole gas? We will proceed this in next section.

2.2 Gibbs Paradox and Entropy

Last chapter we mentioned that the formula of the partition function of the whole gas consisting of N molecules $Z_{whole} = Z^N$ is not correct. That's because we wrongly assume that the identical particles are distinguishable. Why this assumption is wrong? Now let's find the answers.

2.2.1 Gibbs Paradox

To illustrate the Gibbs paradox, let's consider a scenario below. Imagine there is a box under consistent temperature, and it is divided in two parts A and B with the same column V by a partition. Then we fill different kinds of gases say X and Y in these two parts, and both of the gases have the same number of molecules, say N . Then we simply pull out the partition. Suppose there are no chemical interactions between the gases, and when reaching the equilibrium, the entropy of the whole system will increase since we have mixed two different kinds of gases.

Actually we can calculate the exact magnitude of the increase of entropy. Suppose we still utilize that $Z_{whole} = Z^N$, then before we pull out the partition, the entropy of X or Y is³

$$\begin{aligned} S &= \frac{\partial}{\partial T} (kT \ln Z^N) \\ &= Nk \frac{\partial}{\partial T} \left[T \ln T^{3/2} + T \ln \left(\frac{2\pi mk}{h^2} \right)^{3/2} + \ln V \right] \\ &= Nk \left[\ln(VT^{3/2}) + \ln \left(\frac{2\pi mk}{h^2} \right)^{3/2} + \frac{3}{2} \right] \end{aligned}$$

After we pull out the partition, in equilibrium the column increase to $2V$ for both X and Y. Thus we can calculate the increase of entropy as

$$\Delta S = \Delta S_X + \Delta S_Y = 2Nk [\ln(2VT^{3/2}) - \ln(VT^{3/2})] = 2Nk \ln 2$$

If X and Y are the same kind of gases, the increase of entropy oughts to be zero. That's because the entropy oughts to be an extensive quantity⁴. However, the mathematical arguments above didn't care whether the gases dealing with are identical or not. All the formulas above are universal. Hence, there exists contradiction.

³As we can see, the entropy is not an extensive quantity. That's the problem.

⁴Or you can comprehend the zero of increase of entropy on this way: Although you pull out the partition, nothing happens and the equilibrium will not be broken since the gases in both of the parts of the box are identical.

This is the Gibbs paradox.

2.2.2 Gibbs Modification for partition functions

How to address the problem of the Gibbs paradox? The key is that we need to modify the formula of the partition function Z_{whole} for the identical particles in order to make the entropy of identical gas become an extensive quantity.

We get the old version $Z_{\text{whole}} = Z^N$ since we assume that the identical molecules are distinguishable. But the truth is that all the identical microscopic particles are essentially indistinguishable. By taking this principle into account, now let's fix our notion of partition for identical particles.

Suppose we label the n -th energy eigenstate of i -th particle as $E_n^{(i)}$. Then the partition function for the whole system is

$$Z_{\text{whole}} = \sum_{l,m,n,\dots} e^{-\beta[E_l^{(1)} + E_m^{(2)} + E_n^{(3)} + \dots]}$$

However, if the N molecules are indistinguishable, all the arrangements of (l, m, n, \dots) are denoted the same microstate of the whole system⁵. Thus, the correct formula of Z ought to be

$$Z_{\text{whole}} = \frac{1}{A_n^n} \sum_{l,m,n,\dots} e^{-\beta[E_l^{(1)} + E_m^{(2)} + E_n^{(3)} + \dots]} = \frac{Z^N}{N!} \quad (2.8)$$

This is the right formula of Z_{whole} .

By the way, we can happily get all the conclusion for thermodynamic limit under this form of partition too, since when doing this in last chapter, we always doing derivative to β or μ on $\ln Z_{\text{whole}}$, therefore we can properly ignore the existence of the modification term $N!$.

2.2.3 Entropy

Then let's calculate the entropy for identical gases

$$\begin{aligned} S &= \frac{\partial}{\partial T} (kT \ln Z^N - kT \ln N!) \\ &= Nk \left[\ln(VT^{3/2}) + \ln\left(\frac{2\pi mk}{h^2}\right)^{3/2} + \frac{3}{2} \right] - k(N \ln N - N) \end{aligned}$$

⁵As a simple example, consider the 2 molecule system, the system's microstate of $E_{\text{total}} = E_1^{(1)} + E_2^{(2)}$ is equivalent to the microstate of $E'_{\text{total}} = E_2^{(1)} + E_1^{(2)}$, so they are not two microstates, but represent the same one microstate of the whole system, thus giving a modifying coefficient $\frac{1}{A_2^2}$ for the whole partition function.

$$= Nk \left[\ln \frac{VT^{3/2}}{N} + \ln \left(\frac{2\pi mk}{h^2} \right)^{3/2} + \frac{5}{2} \right] \quad (2.9)$$

where we have utilized the *Stirling formula* to approximate the term $\ln N!$. You can find the Stirling formula in appendix B. This time, the entropy for identical gas is an extensive quantity.

Besides, we can also check that the increase of entropy for the initial scenario. If X and Y are identical gases, the the correct calculation of increase of entropy is

$$\begin{aligned} \Delta S &= S'_{X+Y} - (S_X + S_Y) \\ &= 2Nk \ln \frac{VT^{3/2}}{N} - Nk \ln \frac{VT^{3/2}}{N} - Nk \ln \frac{VT^{3/2}}{N} \\ &= 0 \end{aligned}$$

which return us the right intuition.

2.3 Pressure and Energy

Our goal in this section is to rededuce the other important macroscopic quantities like pressure and energy of the ideal classical gases by using the partition function. This is a brand new way for us to get these important relations, from where we can have an insight about the elegance of statistical mechanics.

2.3.1 Pressure

Now, let's consider a scenario for a closed system⁶ : there is a box contains ideal classical gas, and the system can have thermal interaction with the environment. This is equivalent to say that the system is in canonical ensemble. Suppose there are N particles in the gas, and the temperature of the environment is T . Then firstly let's calculate the pressure of the gas.

According to (1.30), we can calculate the pressure by using the partition function as

$$p = kT \frac{\partial}{\partial V} \ln \frac{Z^N}{N!} = NkT \frac{\partial}{\partial V} \ln Z = NkT \frac{\partial}{\partial V} [\ln V + \ln f(T)] = \frac{NkT}{V}$$

where $f(T)$ is a function of T . Slightly modifying tells us that for ideal classical gas, we have the relation that

$$pV = NkT = \frac{N}{N_A} kN_A T = \nu RT \tag{2.10}$$

where N_A is the *Avogadro constant*, and ν is the *amount of substance*. $R = N_A k$ is called the *universal gas constant*. The equation above is just the famous *state equation of ideal classical gas*.

We've met the state equation of ideal classical gas since we were very young. At that time, we deduced or introduced this formula by using various methods. Now we have managed to deduced the same formula with help of our best friend —partition function.

⁶ Again, since we have the thermodynamic limit, all the results we get below can also be valid for another kinds of systems.

2.3.2 Energy

We can also calculate the (average) energy of the ideal classical gas by using the partition function as

$$\begin{aligned} E &= -\frac{\partial}{\partial \beta} \ln \frac{Z^N}{N!} = NkT^2 \frac{\partial}{\partial T} \ln Z \\ &= NkT^2 \frac{\partial}{\partial T} \left[\ln T^{\frac{t+r+2v}{2}} + \ln g(V) \right] \\ &= \frac{t+r+2v}{2} NkT \end{aligned} \tag{2.11}$$

where t, r, v are the degrees of freedom of translation, rotation and vibration respectively, and $g(V)$ is a function of V .

Equal Division Theorem of Energy

What's more, we can also figure out the average energy for a single molecule as

$$\bar{\epsilon} = \frac{E}{N} = (t + r + 2v) \frac{kT}{2} \tag{2.12}$$

The formula above is named as *the equal division theorem of energy*, which reveals that *for a molecule in ideal classical gas, the average energy is equally divided for each degree of freedom by $\frac{1}{2}kT$, except for the vibration with a different energy of kT* .

Heat Capacity

There are two common kinds of heat capacity. One kind is the heat capacity under consistent volume $C_V = (\frac{\partial E}{\partial T})_V$, another kind is the heat capacity under consistent pressure $C_p = (\frac{\partial H}{\partial T})_p$

The formula of C_V and C_p are easy to get, since we can just use (2.11) and get the heat capacities of the ideal classical gas as

$$C_V = \frac{t+s+2v}{2} Nk = \frac{t+s+2v}{2} \nu R \tag{2.13}$$

and

$$C_p = \frac{\partial(E + pV)}{\partial T} = C_V + \nu R \tag{2.14}$$

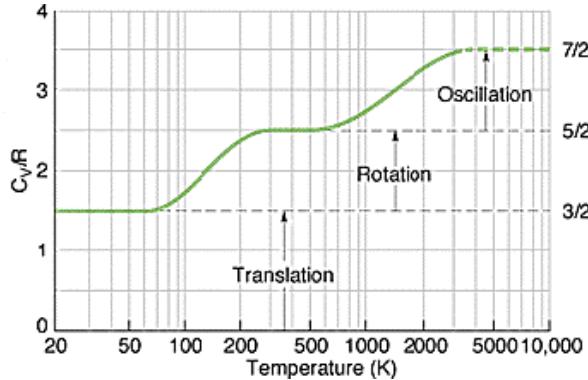


Figure 2.1: The C_V of H_2 under different temperatures. The graph is from David Tong's lecture notes on *Statistical Physics*.

Freeze of the Degrees of Freedom

Armed with the conclusion above, we can know that for diatomic gas, its C_V oughts to be $\frac{7}{2}kT$. But the experimental results for C_V is confusing.

The figure (2.1) shows the C_V of H_2 changing with temperature T from experiments. As we can see, the results are not what we are expected for. We account for the experimental results that under low temperatures, the degrees of freedom of rotation and vibration are frozen. As temperature goes higher, the degrees of freedom of rotation and vibration are excited respectively. As for the reason of this kind of freeze, classical theory has no way out. To properly illustrate this phenomenon, we will have to ask the quantum theory for help. We will revisit this problem in next chapter.

2.4 Maxwell Distribution

For ideal classical gas, the translation kinetic energy of one of its molecules is

$$E = \frac{\mathbf{p}^2}{2m} = \frac{1}{2}mv^2$$

Since there is a probability distribution of E , so there also might be a probability distribution of \mathbf{v} . Our goal in this section is to find this kind of distribution.

2.4.1 Maxwell Distribution of Velocity

Suppose the probability density of \mathbf{p} is $f(\mathbf{p})$, and below we use canonical ensemble to deduce. Thus we can transfer the quantum form of Boltzmann distribution of canonical ensemble into the classical form. The probability to find the velocity located in the tiny phase cell nearby $\frac{\mathbf{p}}{m}$ is

$$\begin{aligned} f(\mathbf{v})d^3\mathbf{v} &= f(\mathbf{p})d^3\mathbf{p} \\ &= \frac{\frac{1}{h^3} \left(\int_{-\infty}^{+\infty} d^3\mathbf{q} \right) e^{-\beta \frac{\mathbf{p}^2}{2m}} d^3\mathbf{p}}{Z} \\ &= \left(\frac{1}{2\pi mkT} \right)^{3/2} e^{-\beta \frac{\mathbf{p}^2}{2m}} d^3\mathbf{p} \\ &= \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-\frac{mv^2}{2kT}} d^3\mathbf{v} \end{aligned}$$

Thus we have the *Maxwell distribution of velocity* as

$$f(\mathbf{v}) = \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-\frac{mv^2}{2kT}} \quad (2.15)$$

Furthermore, if we divide $f(\mathbf{v})$ by

$$f(\mathbf{v}) = f_x(v_x)f_y(v_y)f_z(v_z)$$

then we can get the distribution of components of velocity as

$$f_x(v_x) = \sqrt{\frac{m}{2\pi kT}} e^{-\frac{mv_x^2}{2kT}} \quad (2.16)$$

2.4.2 Maxwell Distribution of Speed

Above we find the probability density of the velocity \mathbf{v} . Now, we want to find the probability density to find a molecule has the speed of v . This, actually, is not

hard.

As long as we have $f(\mathbf{v})$, then let's consider the velocity space, with the coordinate (v_x, v_y, v_z) . Then it's easy to get that

$$f(v)dv = f(\mathbf{v}) 4\pi v^2 dv$$

because the point with the same speed v are distributed on a sphere with radius of v in the velocity space. Thus we ultimately get the *Maxwell distribution of speed*

$$f(v) = 4\pi v^2 f(\mathbf{v}) = 4\pi v^2 \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-\frac{mv^2}{2kT}} \quad (2.17)$$

Then we can calculate lots of things associated with the translation kinetic energy by using the probability densities above, just like when we were learning the thermodynamics and first met the Maxwell distribution. During that time, we deduce the Maxwell distribution by a totally different way.

Chapter 3

Quantum Statistics I

3.1 Energy of Ideal Gases Revisited

Our task in this section is to figure out why the degree of freedom can be frozen for the ideal classical gases. In last chapter, we've got the partition function for classical world. The key to get that is we have

$$e^{\beta \hat{H}} = e^{\beta \hat{T}(\hat{\mathbf{p}})} e^{\beta V(\hat{\mathbf{q}})} + \beta o(\hbar)$$

And then we simply ignore the last term $\frac{o(\hbar)}{kT}$ since we say in classical world \hbar is ignorable. However, there might be a trouble that we haven't noticed before. If the temperature T is very low, then we can't simply ignore the last term anymore. Therefore we can't get the formula (2.2) at low temperature. This argument gives a qualitative explanation for the freeze of degrees of freedom.

To get more precise results, I have to say, in low temperature, we can't use classical things, *thus we'll have to turn to the quantum world*. By doing so, we can naturally find the clues for the freeze of degree of freedom.

3.1.1 Translation

To get the most rigors results of the energy of ideal gases, we need to use the conclusion of quemtum mechanics. The translation kinetic energy of a free quantum particle is

$$E_{\text{trans}} = (n_x^2 + n_y^2 + n_z^2) \frac{\pi^2 \hbar^2}{2mL^2} = (n_x^2 + n_y^2 + n_z^2) E_0$$

Therefore, the partition function of translation is

$$Z_{\text{trans}} = \sum_{n_x} e^{-n_x^2 \beta E_0} \sum_{n_y} e^{-n_y^2 \beta E_0} \sum_{n_z} e^{-n_z^2 \beta E_0}$$

If the temperature is high, i.e. $T \gg \frac{E_0}{k}$, then we can properly approximate this summation by integral as

$$\begin{aligned} Z_{\text{trans}} &= \iiint_0^{+\infty} e^{-(x^2+y^2+z^2)\beta E_0} dx dy dz \\ &= \int_0^{\frac{\pi}{2}} d\phi \int_0^{\frac{\pi}{2}} \sin \theta d\theta \int_0^{+\infty} r^2 e^{-\beta E_0 r^2} dr \\ &= \frac{\pi}{2} \int_0^{+\infty} e^{-\beta E_0 r^2} r^2 dr \\ &= \frac{\pi}{2} \cdot \frac{1}{2\beta E_0} \sqrt{\frac{\pi}{\beta E_0}} \\ &= (2\pi mkT)^{\frac{3}{2}} \frac{2V}{h^3} \end{aligned}$$

which just has a more coefficient 2 compared to what we get in last chapter for Z_{trans} . So we can deduce the same results as classical methods.

If temperature $T \ll \frac{E_0}{k}$ is low, i.e. $\beta E_0 \gg 1$, then the partition function is dominated by the first term. Thus

$$Z_{\text{trans}} \sim e^{-3\beta E_0}$$

So we can calculate that

$$E_{\text{trans}} = -\frac{\partial}{\partial \beta} \ln Z = 3E_0$$

which is to say the particle is on the basic state of translation. However, the order of magnitudes of $\frac{E_0}{k}$ is $10^{-18} K$, thus in most of the case, we can happily use the classical form of the partition function, and the degree of freedom of translation is really hard to be frozen.

3.1.2 Rotation

Then let's talk about rotation. We'll only talk about the diatomic molecules. In the center-of-mass frame, the behavior of the diatomic molecule is just like the hydrogen atom with a proton and an electron. Therefore, according to quantum mechanics, for the degree of freedom of rotation the form of its energy ought to be

$$E_j = \frac{L_j^2}{2I} = \frac{j(j+1)\hbar^2}{2I}$$

where $j = 0, 1, 2 \dots$ is the angular quantum number. What's more, the degeneracy of each E_j is $2j + 1$. So the partition function of rotation is

$$Z_{\text{rot}} = \sum_j (2j + 1) e^{-\frac{\beta \hbar^2}{2I} j(j+1)}$$

If $T \gg \frac{\hbar^2}{2kI}$, then we can properly approximate this summation by integral as

$$\begin{aligned} Z_{\text{rot}} &= \int_0^{+\infty} (2x + 1) e^{-\frac{\beta \hbar^2}{2I} x(x+1)} dx \\ &= \int_0^{+\infty} e^{-\frac{\beta \hbar^2}{2I} (x^2 + x)} d(x^2 + x) \\ &= \frac{2IkT}{\hbar^2} = \frac{8\pi^2 IkT}{h^2} \end{aligned}$$

which is the same as what we get before. So we can deduce the same results as classical methods.

If $T \ll \frac{\hbar^2}{2kI}$, then $Z_{\text{rot}} \sim 1$, thus it has no contributions to the energy. We say that it's frozen.

3.1.3 Vibration

We still use the same scenario to describe a chemical bond, which is a harmonic oscillator with frequency w connecting atoms A and B . According to quantum mechanics, the energy level for one dimensional harmonic oscillator is

$$E_n = \left(\frac{1}{2} + n \right) \hbar w$$

with no degeneracy. Thus, we can write down the partition function directly as

$$\begin{aligned} Z_{\text{vib}} &= \sum_n e^{-\beta(\frac{1}{2} + n)\hbar w} \\ &= e^{-\frac{\beta \hbar w}{2}} \sum_n (e^{-\beta \hbar w})^n \\ &= \frac{e^{-\beta \hbar w/2}}{1 - e^{-\beta \hbar w}} \\ &= \frac{1}{2 \sinh(\beta \hbar w/2)} \end{aligned}$$

3.1. ENERGY OF IDEAL GASES REVISITED

At high temperature $T \gg \frac{\hbar w}{k}$, since we have $\sinh x \sim x$ when $|x| \ll 1$, thus the partition function becomes

$$Z_{\text{vib}} = \frac{kT}{\hbar w}$$

which is exactly the same as what we got before by using classical methods.

If $T \ll \frac{\hbar w}{k}$, then the vibration partition function is dominated by its first term, thus becoming

$$Z_{\text{vib}} = e^{-\frac{\beta \hbar w}{2}}$$

Therefore we can calculate that

$$E_{\text{vib}} = -\frac{\partial}{\partial \beta} \ln Z = \frac{1}{2} \hbar w$$

which simply means that the oscillator is keeping on the basic state, and can't reach the higher energy levels. Thus under low temperature, the vibration modes have no contributions to the heat capacity. It is frozen.

3.2 Photon Gases

Consider a box under certain temperature, on the one hand, the molecules consisting the wall of the box will emit electromagnetic radiation into the cavity, on the other hand, the wall will also absorb the radiation in the cavity. These two process will ultimately reach a equilibrium. In equilibrium, it can be viewed that the cavity is filled with the "photon gas".

What we want to research this section is photon gas, or in other words the electromagnetic radiation, contained in a container in thermal equilibrium. It turns out that the photon gas is a system that is different from the other normal systems that we meet in statistical mechanics because the number of photons in an isolated system is not conserved.

3.2.1 Density of States in 3D box

Initially, let's deduce a useful formula of the density of states of the photon gas. Consider a free relativistic quantum particle confined in a 3D box with volume of $V = L^3$. According to the conclusion of relativistic quantum mechanics, the energy

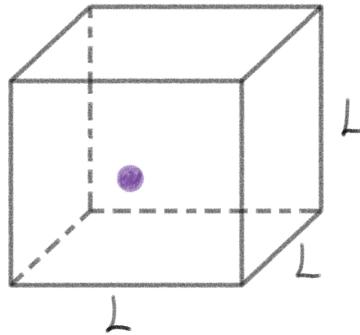


Figure 3.1: A free relativistic quantum particle in 3D box

level of this particle is

$$E = \sqrt{p^2 c^2 + m^2 c^4} = \sqrt{(n_x^2 + n_y^2 + n_z^2) \frac{\pi^2 \hbar^2 c^2}{L^2} + m^2 c^4} \quad (3.1)$$

and we still use (n_x, n_y, n_z) to label a microstate. When E is very large, the number of the microstates with the same energy E can be ridiculously large. Now let's find the density of states for this particle when E is large.

We denote the *density of states* as $g(E)$ so that $g(E)dE$ equals to the number of

microstates that have the magnitude of energy located in the infinitesimal interval $[E, E+dE]$. To achieve this, firstly let's consider the total number of microstates with energy lower than E , which is denoted as $\Phi(E)$. If we can find the formula of $\Phi(E)$, then we can get $g(E)$ by $g(E)dE = \Phi(E + dE) - \Phi(E) = d\Phi$, i.e. $g(E) = \Phi'(E)$.

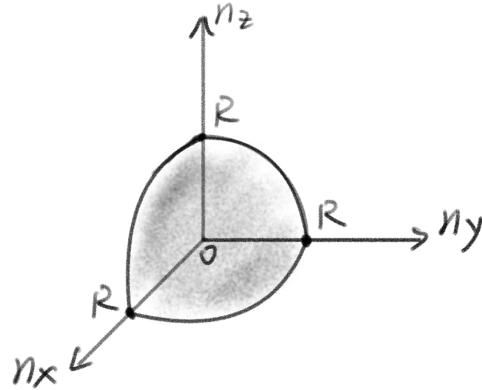


Figure 3.2: The space spanned by n_x, n_y and n_z .

To get $\Phi(E)$, we can consider a space whose 3 axes are n_x, n_y and n_z . Then we denote $R^2 = n_x^2 + n_y^2 + n_z^2$. Then the microstates with energy E simply distribute on the $\frac{1}{8}$ sphere with the radius of

$$R = \sqrt{n_x^2 + n_y^2 + n_z^2} = \sqrt{\frac{L^2}{\pi^2 \hbar^2 c^2} (E^2 - m^2 c^4)}$$

It's evident that $\Phi(E)$ equals to the $\frac{1}{8}$ volume of this ball because the volume cell occupied by one microstate equals to $dn_x dn_y dn_z = 1 \times 1 \times 1 = 1$. Thus, we can get that

$$\begin{aligned} \Phi(E) &= \frac{1}{8} \cdot \frac{4}{3} \pi R^3 \\ &= \frac{\pi}{6} \left[\frac{L^2}{\pi^2 \hbar^2 c^2} (E^2 - m^2 c^4) \right]^{3/2} \\ &= \frac{V}{6\pi^2 \hbar^3 c^3} (E^2 - m^2 c^4)^{3/2} \end{aligned}$$

Thus, the density of states is

$$g(E) = \frac{VE}{2\pi^2 \hbar^3 c^3} \sqrt{E^2 - m^2 c^4} \quad (3.2)$$

Particularly, for photons that we are interested in this section, they can be seen as massless free relativistic quantum particles. And since there are two different polarization states for photons with frequency of ν , so there ought to have a coefficient 2. Thus

$$g(E) = \frac{VE^2}{\pi^2\hbar^3c^3} \quad (3.3)$$

By the way, why we only take 1/8 of the ball into account? That's because the microstates of (n_x, n_y, n_z) and $(-n_x, n_y, n_z)$ are equivalent in the sense of quantum mechanics.

Considering that for a photon

$$E = h\nu = \hbar\omega \quad (3.4)$$

so we can also find the density of states as a function of ν or ω . Since

$$g(\nu)d\nu = g(E)dE = \frac{Vh^2\nu^2}{\pi^2\hbar^3c^3}d(h\nu) = \frac{8\pi V}{c^3}\nu^2d\nu$$

and

$$g(\omega)d\omega = g(E)dE = \frac{V(\hbar\omega)^2}{\pi^2\hbar^3c^3}d(\hbar\omega) = \frac{V}{\pi^2c^3}\omega^2d\omega$$

Thus we can get that

$$\begin{cases} g(\nu) = \frac{8\pi V}{c^3}\nu^2 \\ g(\omega) = \frac{V}{\pi^2c^3}\omega^2 \end{cases} \quad (3.5)$$

We will immediately use the (3.5) below when finding the partition function of photon gas.

3.2.2 Plank Distribution and Blackbody Radiation

Now we want to calculate the energy of the photon gas. We need to figure out the partition function first.

This time, we can't research the photon one by one, since the photon might be absorbed or emitted by the wall of the 3D box. Thus, the number of photon is not conserved for the combined isolated system. Therefore, we will research the whole photon gas directly.

As mentioned above, the photon gas is contained by a 3D box, and the wall of box can absorb and emit the photons simultaneously. In equilibrium, the photon gas and the box will have the same temperature, the definition of which is still $\frac{1}{T} = \frac{\partial S}{\partial E}$. We view the box as a reservoir with consistent temperature T . Since the number of

photon N is changeable, and the N_{total} of the combined system is not conserved, so we reasonably assume that the entropy of photon gas must be independent of N . We view the entropy of photon gas as a function of E and V . Therefore, by using the same method in chapter 1, we can analyze that for photon gas contained in a box with consistent volume, its partition function has the same form as that of canonical ensemble

$$Z = \sum_n e^{-\beta E_n}$$

where the summation is over all the microstates of the whole photon gas.

Photon Modes

Now the problem is that *how to determine a microstate and the corresponding energy for photon gas?* The determination of microstates of photon gas is a little different from the systems consisting of normal quantum particles. To address this, we need the concept of *photon modes*.

In an electromagnetic radiation field, we say a certain wave vector \mathbf{k} and polarization mode s label a certain photon mode (\mathbf{k}, s) . The photons with the same magnitude of \mathbf{k} may have different directions of it, thus having different modes. According to the results in last subsection, the number of modes with frequency of $\nu = \frac{ck}{2\pi}$ is $g(\nu)d\nu$, which have already taken the polarizations into account. Different modes of photon are independent, therefore *we can identify a microstate of the photon gas by identifying how many photons are staying on each mode*, and further write down its partition function.

Partition Functions of Photon Gases

Armed with the concept of photon modes, the partition function of the photon gas can be written as

$$Z = \prod_\nu (Z_\nu)^{g(\nu)d\nu} \quad (3.6)$$

where the Z_ν is the partition function of each mode with the same frequency of ν . Since the number of photons staying in any one mode is arbitrary, thus the achievable energy for one mode with frequency ω is $\{0, h\nu, 2h\nu, \dots\}$ ¹. So the formula of Z_ν is

$$Z_\nu = \sum_n e^{-\beta nh\nu} = \frac{1}{1 - e^{-\beta h\nu}} \quad (3.7)$$

¹Ah, although we have used lots of modern concepts up to now, but this set of discrete achievable energy exactly reveals the birth of energy quantam! Isn't it?

Thus the whole partition function satisfies that

$$\ln Z = \int_0^{+\infty} g(\nu) d\nu \ln Z_\nu = -\frac{8\pi V}{c^3} \int_0^{+\infty} \ln(1 - e^{-\beta h\nu}) \nu^2 d\nu \quad (3.8)$$

By the way, it's not hard to understand that (3.7) is built on the assumption that all the identical photon are indistinguishable, which avoids the Gibbs paradox.

Plank Distribution

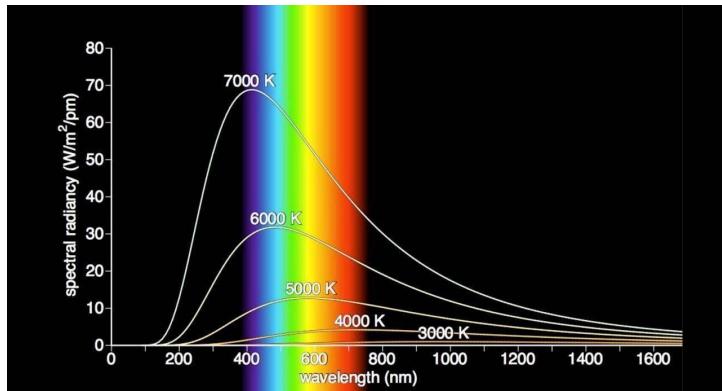


Figure 3.3: Plank distribution

Now we can calculate the (average)² energy of the photon gas as

$$E = -\frac{\partial \ln Z}{\partial \beta} = \frac{8\pi V}{c^3} \int_0^{+\infty} \frac{h\nu^3}{e^{\beta h\nu} - 1} d\nu$$

We define the *energy density of photon gas* $u(\nu, T)$, and $ud\nu$ is the energy per volume of photon gas with frequency between $[\nu, \nu + d\nu]$. Thus we have

$$u(\nu, T) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{kT}} - 1} \quad (3.9)$$

(3.9) is the famous formula of *Plank distribution*. It describes the energy distribution with frequency under certain temperature of the photon gas in equilibrium. We can also convert the formula (3.9) to its version of wavelength as

$$u(\lambda, T) = \frac{8\pi h}{\lambda^3} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1} \quad (3.10)$$

²You know why I add a bracket here ...

The graph of (3.10) can be seen in figure (3.3).

The (3.9) and (3.10) are the energy densities that we deduced from the photon gases contained in a box. However, it's obvious that for any container with arbitrary shape, they will always hold since they are the energy per volume.

Blackbody Radiation

Photons are also electromagnetic radiation, so the (3.9) is also called the *energy density of electromagnetic radiation*. Actually, it have another famous name, the *Plank formula of blackbody radiation*. What is blackbody radiation? Why the photon gas in equilibrium is concerned with the blackbody radiation?

For an object, on the one hand it can absorb as well as reflect the radiation, on the other hand it can emit its own radiation. Blackbody is the object that can absorb all the radiation without reflecting. If you put a blackbody into a radiation field, the absorption and emission will ultimately reach equilibrium.

For a blackbody put in radiation field in equilibrium, it will have the same temperature as the radiation field, i.e. the photon gas around it. Equilibrium also means that the photon energy power absorbed and emitted by unit area on the surface of the blackbody are the same, regardless of the shape of the blackbody.

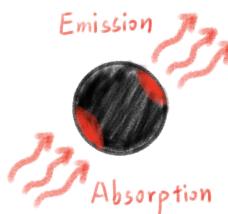


Figure 3.4: A blackbody put in the radiation field in equilibrium

Thus, suppose we have a box containing photon gases with temperature T in equilibrium, then we open a little window with area of S . According to the arguments above, the radiation emitted from the window of the box will be equivalent to the radiation emitted by the surface of a blackbody with same area of S in the same temperature. So that's why we also call the photon gases the blackbody radiation.

By the way, up to now we have discussed *the first dark cloud* covering the sky of physics in 20th century.



Figure 3.5: "Blackbody" radiation

Stefan-Boltzmann Law

Finally, let's deduce the radiation power emitted by unit area of the cover of a blackbody, which is the same as the flux density of radiation.

According to the arguments above, we can identify blackbody radiation with the photon gas. So let's firstly consider a box filled with photon gas in temperature T . Then we open a window with area of S , and calculate the power of radiation emitted from this window.

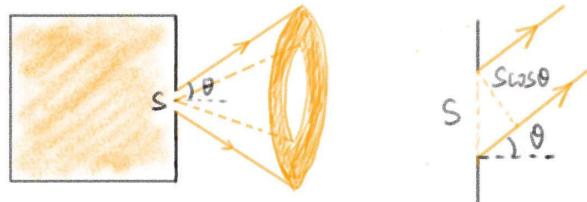


Figure 3.6: Left: The radiation in the solid angular $d\Omega$ near by the direction of θ ; Right: The reason why you should multiply a $\cos\theta$.

The energy pass through the window along the solid angle $d\Omega$ near the direction of θ in time interval Δt is

$$\Delta E = \frac{d\Omega}{4\pi} \xi \cdot c \Delta t S \cos \theta$$

where c is the speed of light and ξ is the integral of u over all frequencies³. Then if we want to calculate the radiation power emitted through the window, we just need to do integral on the solid angular. Thus

$$P = \int \frac{\Delta E}{\Delta t} = \frac{\xi c S}{4\pi} \int \cos \theta d\Omega = \frac{\xi c S}{2} \int_0^{\pi/2} \sin \theta \cos \theta d\theta = \frac{1}{4} \xi c S$$

³ ξ is just the energy of the photon gas in unit volum. Thus the term $\frac{d\Omega}{4\pi} \xi$ is the energy in unit volum that emitted along the solid angular $d\Omega$. This holds since the photon gas is isotropous.

We further define the *flux density of radiation* as

$$J = \frac{P}{S} = \frac{1}{4}\xi c$$

which describes the energy emitted by the unit area of the cover of a blackbody in unit time period. Find the exact formula of J is our ultimate goal.

The last work is to determine ξ . According to its definition, we can get that⁴

$$\xi = \int_0^{+\infty} u(\nu, T) d\nu = \frac{8\pi h}{c^3} \int_0^{+\infty} \frac{\nu^3 d\nu}{e^{\frac{h\nu}{kT}} - 1} = \left(\frac{8\pi k^4}{c^3 h^3} \int_0^{+\infty} \frac{x^3 dx}{e^x - 1} \right) T^4 = \frac{\pi^2 k^4}{15 c^3 h^3} T^4$$

Thus we can ultimately get the *Stefan-Boltzmann law* of blackbody radiation

$$J = \sigma T^4 \tag{3.11}$$

where $\sigma = \frac{\pi^2 k^4}{60 c^2 h^3}$ is the *Stefan-Boltzmann constant*. In history, Stefan found that J is proportional to T^4 and Boltzmann deduced this relation by using the statistical methods.

3.2.3 Radiation Pressure

It's not hard to deduce the pressure of the ideal photon gas in equilibrium, or equivalently the blackbody radiation pressure in equilibrium, by using the partition function that we got initially in this section

$$\begin{aligned} p &= \frac{1}{\beta} \frac{\partial \ln Z}{\partial V} \\ &= -\frac{8\pi}{\beta c^3} \int_0^{+\infty} \ln(1 - e^{-\beta h\nu}) \nu^2 d\nu \end{aligned}$$

⁴The tricky integral can be dealt as below

$$\begin{aligned} \int_0^{+\infty} \frac{x^3 dx}{e^x - 1} &= \int_0^{+\infty} x^3 \frac{e^{-x}}{1 - e^{-x}} dx = \int_0^{+\infty} x^3 \sum_{n=1}^{\infty} e^{-nx} dx \\ &= \sum_{n=1}^{\infty} \int_0^{+\infty} x^3 e^{-nx} dx = \sum_{n=1}^{\infty} \frac{1}{n^4} \int_0^{+\infty} (nx)^3 e^{-nx} d(nx) \\ &= \left(\int_0^{+\infty} t^{4-1} e^{-t} dt \right) \cdot \sum_{n=1}^{\infty} \left(\frac{1}{n} \right)^4 \\ &= \Gamma(4) \cdot \zeta(4) \\ &= \frac{\pi^4}{15} \end{aligned}$$

where Γ is the gamma function and ζ is the Riemann zeta function. Thus, as a physicist, it can be a wonderful thing if you have a friend who is a mathematician.

$$\begin{aligned}
 &= -\frac{8\pi}{\beta c^2} \int_0^{+\infty} \ln(1 - e^{-\beta h\nu}) d\left(\frac{\nu^3}{3}\right) \\
 &= \frac{8\pi}{\beta c^3} \int_0^{+\infty} \frac{\nu^3}{3} d\ln(1 - e^{-\beta h\nu}) - \frac{8\pi}{\beta c^3} \left[\frac{\nu^3}{3} \ln(1 - e^{-\beta h\nu}) \right] \Big|_0^{+\infty} \\
 &= \frac{8\pi h}{3c^3} \int_0^{+\infty} \frac{\nu^3 e^{-\beta h\nu}}{1 - e^{-\beta h\nu}} d\nu + 0 \\
 &= \frac{8\pi k^4 T^4}{3c^3 h^3} \int_0^{+\infty} \frac{x^3}{e^x - 1} dx \\
 &= \frac{\pi^2 k^4 T^4}{45c^3 h^3}
 \end{aligned}$$

By citing the formula of the Stefan-Boltzmann constant, we can write the radiation pressure of photon gas as

$$p = \frac{4\sigma}{3c} T^4 \quad (3.12)$$

This kind of radiation pressure is essential in fighting against the gravity thus maintaining the stability of a star's structure.

3.3 Heat Capacity of Solid

Now let's discuss *the second dark cloud* covering the sky of physics in 20th century.

3.3.1 Dulong - Petit Law

First of all, let's calculate the heat capacity of elemental crystal by using classical methods.

For a crystal consists of N particles, since different particles are in different equilibrium position. so they are distinguishable.

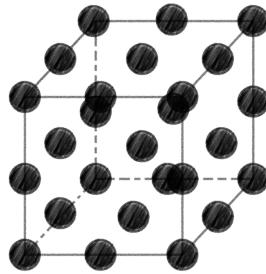


Figure 3.7: A crystal

Each particle can vibrate along 3 independent directions x, y and z . Thus the degree of freedom for the whole crystal is $3N$, and they are all degree of freedom of vibration.

According to the equal division theorem of energy, the energy of this crystal is.

$$E = 3NkT = 3\nu RT \quad (3.13)$$

Thus, the heat capacity per mole under consistent pressure is

$$C_V^{\text{mol}} = \frac{1}{\nu} \frac{\partial E}{\partial T} = 3R \quad (3.14)$$

It's a constant for elemental crystal, which is called the *Dulong - Petit law*.

However, experiments have found the violation of this law under low temperature, since all the capacities of solid decrease to zero as temperature goes low. This problem is quite similar to the freeze of degree of freedom of ideal gas under low temperature. So the methods to address them can also be similar —we have to ask quantum for help.

3.3.2 Einstein's Model

In history, Einstein first managed to explain the violation of Dulong - Petit law by using the concept of quantum, which was seen us a successful application of the early quantum theory.

The idea of *Einstein's model* is simply that we view each degree of freedom of vibration as a quantum harmonic oscillator with some frequency of ω . So the energy levels are

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega, n = 0, 1, 2 \dots$$

without degeneracy. Thus the partition function for a single degree of freedom of vibration is

$$z = \sum_n e^{-\beta(n+\frac{1}{2})\hbar\omega} = e^{-\frac{\beta\hbar\omega}{2}} \sum_n e^{-n\beta\hbar\omega} = \frac{e^{-\frac{\beta\hbar\omega}{2}}}{1 - e^{-\beta\hbar\omega}}$$

Now that the N particles are distinguishable, thus the whole partition function is

$$Z_{\text{whole}} = z^{3N}$$

Thus the energy of the crystal is

$$\begin{aligned} E &= -\frac{\partial}{\partial\beta} \ln z^{3N} \\ &= 3N \frac{\partial}{\partial\beta} \left[\ln(1 - e^{-\beta\hbar\omega}) + \frac{\beta\hbar\omega}{2} \right] \\ &= 3N \frac{\hbar\omega e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}} + \frac{1}{2} \hbar\omega \\ &= \frac{3N}{2} \hbar\omega + \frac{3N\hbar\omega}{e^{\frac{\hbar\omega}{kT}} - 1} \end{aligned}$$

Thus, the C_V^{mol} is

$$C_V^{\text{mol}} = \frac{1}{\nu} \frac{\partial E}{\partial T} = 3R \left(\frac{\hbar\omega}{kT} \right)^2 \frac{e^{\frac{\hbar\omega}{kT}}}{\left(e^{\frac{\hbar\omega}{kT}} - 1 \right)^2} \quad (3.15)$$

It can be analyzed that at high temperature $T \gg \frac{\hbar\omega}{k}$, we have $\frac{\hbar\omega}{kT} \ll 1$, thus we have

$$C_V^{\text{mol}} \approx 3R \lim_{x \rightarrow 0} \frac{x^2 e^x}{(e^x - 1)^2} = 3R \lim_{x \rightarrow 0} \frac{x^2(x+1)}{(x+1-1)^2} = 3R$$

which returns back to the Dulong - Petit law.

At low temperature $T \ll \frac{\hbar\omega}{k}$, $e^{\frac{\hbar\omega}{kT}} - 1 \approx e^{\frac{\hbar\omega}{kT}}$, so we have

$$C_V^{\text{mol}} \approx 3R \left(\frac{\hbar\omega}{kT} \right)^2 e^{-\frac{\hbar\omega}{kT}} \quad (3.16)$$

Since it has the factor $e^{-\frac{\hbar\omega}{kT}}$, C_V^{mol} drops too fast to zero as temperature goes lower, which violates the experiments. But, qualitatively, the Einstein's model is the first to provide the explanation for why C_V^{mol} of the solid decrease to zero at low temperature.

3.3.3 Debye's Model

Phonon

Einstein's model research the particles' vibration in a crystal, and we need to identify the microstates of the crystal by identifying the microstates of all the particles in it.

However, there is another model named *Debye's model*, which provides a brand new viewpoint. Debye's model view the solid as a continuous medium, and similar to electromagnetic field, there exists the vibration field in the solid. The wave corresponding to this kind of vibration is just like sound wave. This kind of vibration field is excited thermally, and has the quantum property too.

As the photon is the particle of electromagnetic field, we call the "particle" of the vibration field *phonon*. Just like the ideal photon gas, you can view a solid as filled with phonon gas. However, the phonon is not the real particle, thus we call it a kind of *quasi - particle*.

The idea is that entropy of the crystal is exactly the entropy of the phonon gas, so as too the energy. Thus, we now turn to research the phonon gas.

Partition Function

The vibration field in crystal has two different types, transverse vibration and longitudinal vibration. Thus, for a phonon with the same wave vector k , it will have 3 different as well as independent modes.

The phonons are of course massless, so we could only treat them as relativistic free quantum "particles". Thus, the density of states of phonon gas is

$$g_s(\nu) = \frac{12\pi V}{c_s^3} \nu^2 \quad (3.17)$$

Compared it with the density of states of photon, firstly the coefficient is transferred

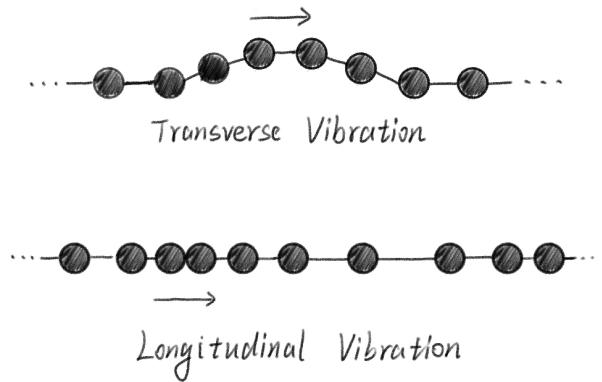


Figure 3.8: Two kinds of vibrations

from 8 to $8 \times \frac{3}{2} = 12$, since we have longitudinal vibration for phonons; Secondly $c_s = \lambda_s \nu$ is the speed of phonons in solid. It's also the speed of sound.

There's also one more difference between phonons and photons. Essentially, phonons are produced by the vibration of particles in the crystal, thus its wavelength will have a lower limit λ_m . Evidently, λ_m is in the same order of magnitude of the distance between the particles in the crystal. Thus, there oughts to be an upper limit for the frequency for a phonon which is denoted as ν_m , while the frequency of a photon has no upper limits in principle.

Now, we can write down the partition function for the phonon gas. Firstly we divide the partition function by

$$Z = \prod_{\nu} Z_{\nu}^{g_s(\nu)d\nu}$$

where

$$Z_{\nu} = \sum_n e^{-\beta n \hbar \nu} = \frac{1}{1 - e^{-\beta \hbar \nu}}$$

Thus we have

$$\ln Z = \int_0^{\nu_m} g_s(\nu) \ln Z_{\nu} d\nu = - \frac{12\pi V}{c_s^3} \int_0^{\nu_m} \ln(1 - e^{-\beta \hbar \nu}) \nu^2 d\nu \quad (3.18)$$

Heat Capacity

Now we can calculate the energy, and further calculate the heat capacity of a solid by using the partition function of phonon.

$$E = -\frac{\partial}{\partial \beta} \ln Z = \frac{12\pi V}{c_s^3} \int_0^{\nu_m} \frac{\hbar\nu^3}{e^{\beta\hbar\nu} - 1} d\nu = \frac{12\pi V \hbar}{c_s^3} \left(\frac{kT}{\hbar}\right)^4 \int_0^{\frac{\hbar\nu_m}{kT}} \frac{x^3}{e^x - 1} dx \quad (3.19)$$

At high temperature $T \gg \frac{\hbar\nu_m}{k}$, $\frac{\hbar\nu_m}{kT} \ll 1$, thus we can reasonably approximate that

$$\begin{aligned} E &= \frac{12\pi V \hbar}{c_s^3} \left(\frac{kT}{\hbar}\right)^4 \int_0^{\frac{\hbar\nu_m}{kT}} \frac{x^3}{e^x - 1} dx \\ &\approx \frac{12\pi V \hbar}{c_s^3} \left(\frac{kT}{\hbar}\right)^4 \int_0^{\frac{\hbar\nu_m}{kT}} x^2 dx \\ &= \frac{4\pi V k \nu_m^3}{c_s^3} T \end{aligned}$$

According to the Dulong - Petit Law, the result under high temperature oughts to be $3NkT$. Thus we can deduce that

$$\nu_m = \left(\frac{3Nc_s^3}{4\pi V}\right)^{1/3} \quad (3.20)$$

At low temperature, $\frac{\hbar\nu_m}{kT} \gg 1$, then we can approximate that

$$\begin{aligned} E &\approx \frac{12\pi V \hbar}{c_s^3} \left(\frac{kT}{\hbar}\right)^4 \int_0^{+\infty} \frac{x^3}{e^x - 1} dx \\ &= \frac{12\pi V \hbar}{c_s^3} \left(\frac{kT}{\hbar}\right)^4 \frac{4\pi^4}{15} \\ &= \frac{4\pi^5 V k^4}{5c_s^3 h^3} T^4 \end{aligned}$$

Therefore, we have the heat capacity under consistent pressure as

$$C_V = \frac{\partial E}{\partial T} = \frac{16\pi^5 V k^4}{5c_s^3 h^3} T^3 = \frac{12\pi k^4 N}{5\nu_m^3 h^3} T^3 \quad (3.21)$$

Thus we have the famous $C_V \propto T^3$ which is compatible finely with the experiments under low temperature as shown in the (3.9). In most materials the heat capacity is dominated by the phonon contribution. (In metals there is an additional contribution from conduction electrons that we will calculate in next chapter.)

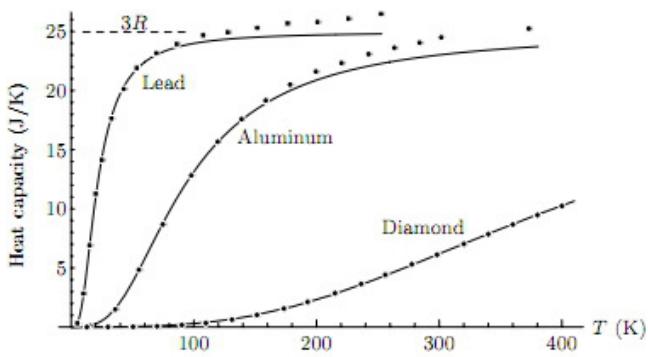


Figure 3.9: Experimental data for heat capacities. The solid line is the Debye prediction. The graph comes from David Tong's lecture notes on *Statistical Physics*.

Chapter 4

Quantum Statistics II

4.1 Distribution of Particle number

4.1.1 Maxwell-Boltzmann Distribution

Now let's discuss a new question : what's the distribution of the number of particles on each energy microstate ?

Since under the thermodynamic limit, microcanonical, canonical and grand canonical ensemble coincide, thus the distributions of them are also the same. For the sake of calculation, we choose the grand canonical ensemble.

Suppose we are talking about the identical particles which are indistinguishable. So the grand canonical partition function is

$$\Xi = \sum_{m,r} \frac{e^{-\beta(E_{r,m} - \mu N_m)}}{N_m!} = \sum_{N_m} e^{\beta \mu N_m} \frac{1}{N_m!} Z^{N_m} = \sum_{N_m} \frac{(Z e^{\beta \mu})^{N_m}}{N_m!} = \exp(e^{\beta \mu} Z)$$

where $Z = \sum_r e^{-\beta E_r}$ is the partition function for a single particle. Then we can calculate the (average) particle number as

$$N = \sum_r n_r = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \Xi = \frac{1}{\beta} \frac{\partial}{\partial \mu} (e^{\beta \mu} Z) = Z e^{\beta \mu} = \sum_r e^{-\beta(E_r - \mu)}$$

where we further denote n_r as the number of particles staying on the microstate of $|E_r\rangle$. Thus we can identify that

$$n_r = e^{-\beta(E_r - \mu)} = \frac{1}{e^{\beta(E_r - \mu)}} \quad (4.1)$$

This is the formula of *Maxwell-Boltzmann distribution*, where β and μ reflect the temperature and chemical potential of the *system*.

4.1.2 Bose-Einstein Distribution

Above we deduce the partition function of a system by first confirming the partition function of a single particle, and then construct the whole partition function.

However, there is another way to think. Suppose that for a system, all the energy states are determined and independent with each other. Thus we can identify the microstate of the system by confirming how many particles are at state 1, 2... and so on. And, for grand canonical ensemble, we assume that the particle number on a certain state can be arbitrary. Thus, we can divide the whole partition function as

$$\Xi = \prod_r \Xi_r$$

where

$$\Xi_r = \sum_N e^{-\beta(NE_r - \mu N)} = \sum_N [e^{-\beta(E_r - \mu)}]^N = \frac{1}{1 - e^{-\beta(E_r - \mu)}}$$

is the partition for a certain energy state $|E_r\rangle$. The answer above holds for all the $\mu < 0$.

Then we can calculate the number of particles by using the grand partition function as

$$N = \sum_r n_r = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \Xi = \sum_r -\frac{1}{\beta} \frac{\partial}{\partial \mu} \ln [1 - e^{-\beta(E_r - \mu)}] = \sum_r \frac{1}{e^{\beta(E_r - \mu)} - 1}$$

Thus we can identify that

$$n_r = \frac{1}{e^{\beta(E_r - \mu)} - 1} \quad (4.2)$$

which is the formula of *Bose-Einstein distribution*.

4.1.3 Fermi-Dirac Distribution

Above we deduce the Bose-Einstein distribution by assuming that the particle number for each energy state can be arbitrary. We call the particles satisfying this property Bosons.

However, there are also another kind of particles which satisfy that each energy state can at most contain one particle. We call this kind of particles Fermions.

Hence, the Ξ_r for Fermions are

$$\Xi_r = e^{-\beta(0 \cdot E_r - \mu \cdot 0)} + e^{-\beta(1 \cdot E_r - \mu \cdot 1)} = 1 + e^{-\beta(E_r - \mu)}$$

So we can calculate the particle number as

$$N = \sum_r n_r = \sum_r \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \Xi_r = \sum_r \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln [1 + e^{-\beta(E_r - \mu)}] = \sum_r \frac{1}{e^{\beta(E_r - \mu)} + 1}$$

Thus we can indentify that

$$n_r = \frac{1}{e^{\beta(E_r - \mu)} + 1} \quad (4.3)$$

This is the formula of *Fermi-Dirac distribution*, which holds for both $\mu > 0$ and $\mu < 0$.

4.1. DISTRIBUTION OF PARTICLE NUMBER

Appendix A

Gauss Integrals

The general form of Gauss integral is that

$$G_n = \int_{-\infty}^{+\infty} x^{2n} e^{-ax^2} dx$$

Now let's find its result step by step. First of all, let's consider the simplest situation $n = 0$. Noticing that the square of G_0 is

$$\begin{aligned} G_0^2 &= \int_{-\infty}^{+\infty} e^{-ax^2} dx \int_{-\infty}^{+\infty} e^{-ay^2} dy \\ &= \iint_{-\infty}^{+\infty} e^{-a(x^2+y^2)} dx dy \\ &= \int_0^{2\pi} d\varphi \int_0^{+\infty} e^{-a\rho^2} \rho d\rho \\ &= \frac{\pi}{a} \int_0^{+\infty} e^{-a\rho^2} d(a\rho^2) \\ &= \frac{\pi}{a} \end{aligned}$$

where we simply ask the polar coordinate system for help. Thus, we have that $G_0 = \sqrt{\frac{\pi}{a}}$. Then, it's easy to know that

$$G_{n+1} = \int_{-\infty}^{+\infty} x^{2n+2} e^{-ax^2} dx = -\frac{d}{da} \int_{-\infty}^{+\infty} x^{2n} e^{-ax^2} dx = -\frac{d}{da} G_n$$

Thus, we have that

$$G_n = (-1)^n \frac{d^n}{da^n} G_0 = \sqrt{\pi} (-1)^n \frac{d^n}{da^n} a^{-\frac{1}{2}} = \frac{(2n-1)!!}{(2a)^n} \sqrt{\frac{\pi}{a}}$$

Appendix B

Stirling Formula

The Stirling formula is that for large N , we have

$$\ln n! \approx n \ln n - n$$

To prove this formula, you just need to do the approximate calculation

$$\ln n! = \sum_{i=1}^n \ln i \approx \int_1^n \ln x dx = n \ln n - n$$

Actually, there's a more accurate form of Stirling formula

$$\ln n! \approx n \ln n - n + \frac{1}{2} \ln(2\pi n)$$

Then let's give a brief proof for it. It can be confirmed that

$$n! = \int_0^\infty e^{-x} x^n dx$$

Suppose there's a function $F(x)$ satisfying that

$$e^{-x} x^n = e^{-F(x)}$$

i.e. $F(x) = x - n \ln x$. Then we do Taylor expansion of $F(x)$ as

$$F(x) = F(x_0) + F'(x_0)(x - x_0) + \frac{F''(x_0)(x - x_0)^2}{2} + \dots$$

where we request that $F'(x_0) = 0$. Simple calculation gives us that

$$F(x) = n - n \ln n + \frac{(x-n)^2}{2n} + \dots$$

Thus, for large n , we have

$$e^{-F(x)} = e^{-\left[n - n \ln n + \frac{(x-n)^2}{2n} + \dots\right]} \approx n^n e^{-n} \cdot e^{-\frac{(x-n)^2}{2n}}$$

Therefore

$$\begin{aligned} n! &= \int_0^\infty e^{-F(x)} \approx n^n e^{-n} \int_0^\infty e^{-\frac{(x-n)^2}{2n}} dx \\ &= n^n e^{-n} \int_{-n}^\infty e^{-\frac{x^2}{2n}} dx \\ &\approx n^n e^{-n} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2n}} dx \\ &= n^n e^{-n} \sqrt{2\pi n} \end{aligned}$$

So we ultimately get the more accurate version of Stirling formula as

$$\ln n! \approx n \ln n - n + \frac{1}{2} \ln(2\pi n)$$

Appendix C

我与热力学二三事

尽管只是流水账而没有什么文学性色彩，但我很喜欢回顾自己在学习过程中的经历。其实啊，在我写的每一份讲义里我都会花时间去记录在我学习过程中的点点滴滴。只有这样，我们才能够为自己的进步而感到切实地快乐！所以，接下来我来讲一讲我和热力学之间的故事。

C.1 入门教程

我在上小学的时候是没有科学课的，真正接触物理学还是要在八年级的时候；而接触热力学就要等到九年级了¹。九年级的人教版物理教材是“九年级全一册”，里面前几章介绍了在今天我认为被划分为热力学的初步内容，包括有分子动理论、内能、比热容²、内燃机。我在开学前的暑假就从表哥那里“继承”来的教材上看到了“内能 = 分子动能 + 分子势能”，这个概念在课本上是没有的。

也是在九年级开学前到暑假，我们学校的老师组织一部分学习提前学习³，老师上课讲到了内燃机的部分，自然就不得不提到热力学定律了。当时应该是提到了热力学第一定律吧，这个我印象到不是很深刻；我印象比较深刻的是热力学第二定律——一方面是我从来没有见过表述这么“含糊”且绕来绕去的自然“定律”；另一方面，暑假的有一天，在准备高中竞赛的表哥和我相聚在奶奶家，那时候我们对物理学的兴趣势头都非常足，他问我最近在学些什么，我说刚开始学电学。他拿来我的九年级教材翻了翻，看到了我在热学那几章的笔记，上面记了有热力学第二定律。他表示非常惊喜：“嗯，还不错！竟然在其中还能发现热力学第二定律！”⁴

初中的热学也就这么多了。不过值得一提的是，我在九年级上学期的时候买来了高中物理课本，当时用的还是人教版老教材，热学的部分是《选修 3-3》，封面是好看的雾凇。有天晚上计算机课，我没有听，而是在机房里面浏览选修 3-3。内容当然是没有记住多少的，毕竟是浏览；只留下了一个印象，那就是高中热学、选修 3-3 真的好难啊！

C.2 新手村

进入高中，我也就进入了物理学习的新手村。但是，在我们省当年（最后一年）老高考的背景下，选修 3-3 和选修 3-4 作为选考内容，在高二下学期才会学到，而且只用选择一个学习就行了。不知道为什么，我们学校甚至我们市统一要求我们选 3-4，期末考试的卷子只有 3-4 的内容，所以 3-3 的热学部分学校是没有讲的。但这并不妨碍我自己去学对吧！

¹ 尽管在七年级的时候我们就学习了温度和温度计。

² 我们的老师当时说：“比热容和电势就是整个初中物理学最难的两个概念了。”

³ 或者说，内卷。

⁴ 不过我们没有就热力学第二定律进行深入讨论，而是讨论了电学，他向我炫耀了“基尔霍夫方程组”的方法解电路问题。他画了一个很复杂的电路，说：“来，你用欧姆定律！来，你用欧姆定律...”

物理化学

首次遇见热力学竟然是在化学里。化学刚入学肯定是要讲“物质的量”、阿伏伽德罗定律了。阿伏伽德罗定律其实可以浓缩为理想气体状态方程 $pV = nRT$ ，我们的化学老师一上来就给我们亮了这个底牌了。她在说这个公式的时候，YC 同学竟还能说出来 R 的数值！大家无不惊骇。不过，这倒是我第一次知道这个公式；我其实一直很纳闷为什么高中阶段不论是化学还是物理都对这个公式避而不谈。

其实我在高中阶段对于热学的思考都是在化学里，我发现不论是基本原理层面还是化学实验的部分，要想真正理解其中的所以然，是需要不少热力学知识的。化学平衡常数的公式为什么那么有美感，其背后的基本理论是什么？当时我就有一个信念，不用问这一定和热力学甚至和熵有直接关系，而且我励志要上大学之后彻底研究透这个问题⁵。

而且，我印象里还有一个问题就是到底该如何理解吉布斯自由能（变） ΔG ？若一个可逆反应正向可以自发进行，说明正向的自由能变为负，那么其逆反应呢？逆反应肯定也可以自发进行，因为毕竟是可逆反应，但是其自由能变是正的啊？该如何解释呢？而且，到底什么是自由能？这个“能”的物理图像是什么？当时的确有像这样的一连串的疑问，但是在当时得不到解答。惭愧的是，这个话题我到现在还没抽出时间来仔细研究。

还有一次，2022 年的夏天，有一天我们学生因为疫情被封到了宿舍，JCH、LYF、WYB 和我就在宿舍里开研讨会——我们每个人挑一个话题来讲。我讲的内容就是那两天写的学习手记，总结了一下如何运用一些简单的热力学知识去理解高中化学知识。我印象中 JCH 给我们讲了他最喜欢的国学大师陈寅恪。

最后再分享一件关于化学和热学的故事。2022 年 8 月底，高三的时候，有一次在“大本”上遇到了一个概念叫“汽化中心”，我花了一个周末的时间仔细研究了一下什么是汽化中心，以及为什么提供汽化中心就可以避免液体暴沸。我写了一篇文章，那篇文章现在还在我的 iPad 里面。这篇文章中我提出了一个很定性、很简化还有点天马行空的模型来解释和讨论上述论述。我甚至还把这篇文章打印了出来，本来想给化学老师看的，但最后我没有给。

总之啊，用热力学的基本原理来解释和理解化学，在高中学习中的案例不胜枚举。这难怪，因为这些内容构成了物理化学这门学科的重要组成部分——化学热力学。上面的这些问题在科学史上虽然早就被人研究过，但是对于当时的我来讲是很有吸引力的。我在高中的时候总想着这些问题在大学里，在我学习了足够的物理学知识之后，都会得到彻底的解决。但现在真的来到了大学这个阶段之后，我才发觉原来我做的（包括这份统计力学讲义）仅仅是梳理这个领域的基本原理，而我上面所提到的

⁵ 这个问题我在 2024 年夏天，大一升大二的暑假探讨了一次。那时候刚浏览了一遍赵凯华老师的新概念物理《热学》，里面有如何利用熵增原理推导化学平衡常数表达式的线索，于是我就写了一篇文章发在了探星阁公众号上，名称叫做《论化学平衡常数》。

但是这件事后来一直被搁置，直到 2025 年秋天，我学习统计力学的时候，参考了汪志诚老师的《热力学·统计物理》，其热力学部分有介绍如何推导化学平衡条件，本质也是用熵增原理，只不过在表面上用到了吉布斯自由能。

那些问题属于应用的范畴。我似乎没有时间去悉数每一个领域中对于基本原理的应用的部分，所以最终也只能是挑选我最感兴趣的几个去花时间梳理梳理了。

热力学

刚才我也有提到，在高中阶段课内我们是不学热学的，我的热学知识来源于我的竞赛学习。在 2021 年 9 月份准备竞赛的时候，我学习了热学的知识，其实主要就是热力学第一定律和热力学循环那一套。虽然我买了沈克琦老师的《高中物理学》全套，但是第二本《热学》我却没有认真研读，只是草草读了读。但是其中有一个东西很神奇我很有印象，就是麦克斯韦分布，我当时觉得这是多么的不可思议，但是现在我已经掌握了两种方法去推导它了。

还有一件比较有意思的事情，就是高三的时候我问了物理老师一个问题，印象中好像类似是这样：“一个试管塞上塞子倒过来放，那么管内气体的压强是不是大气压强？”我的物理老师说他好久不学热学了都忘了，还说这个问题听起来不是很难的问题。他问我为什么要思考热学的问题高考又不考，我说因为在化学题里面遇到了这个问题。但是他当时忙着要去开会，所以这个问题最后还是我自己去思考清楚了。

最后我想说，前面我提到初三的时候我觉得高中的热学特难，但是上了高中之后，慢慢地我感觉也没啥难的，毕竟是放在了选考的部分。但是物理所的曹则贤老师 2023 年在中科院做了跨年演讲，主题就是热力学。他曾经在微信公众号留言区吐槽说：“看不懂就对啦！热力学其实是最难的，我们的课本上什么都没有讲！”

C.3 打副本

终于盼着上了大学了。有意思的是，大学物理不同专业课的学习就像是打不同的副本一样：不同的课程彼此有比较强的独立性，而同一门课程还有 1.0 和 2.0 版本。2024 年秋冬，大二上学期，我在打热学副本的 1.0 版本；而 2025 年秋冬，我在打其 2.0 版本，也就是统计力学部分。

热学

尽管是 1.0 部分的普物热学也是放在大二学习的。2024 年夏天，我就已经迫不及待地想要开始学习热学了，一方面是因为高中的时候的确对这块思考得相对少，学科框架搭建的不如力学、电磁学；另一方面当时看目录我就发现赵凯华老师的新概念热学里有些我比较感兴趣的话题——晶体（结构化学）、高三的时候就了解到的布朗运动以及爱因斯坦的相关工作以及上面提到的化学中遇到的种种热力学问题等；此外还有因为新概念热学真的相对而言非常薄，所以我有野心能够尽快开始学完——所幸的是，尽管暑假学习效率真的很低，我还是浏览了一遍赵老师的新概念热学。但是这本书真不愧大家都说是热学里最难的，最为课外书、参考书看看还行，因为（站

在中国国内教学大纲的角度说) 其基本的知识讲得不透, 而补充的东西又太多且无法深入。

说实话其实我当时还是很盼望着学习热学的, 因为有好多我之前就有所耳闻却没有深入学的东西比如麦克斯韦分布律、自由能、平均自由程等等, 想着在学习热学的时候认真学学。但是说实话, 普物热学的学习也没能帮我实现梦想。因为当时普物的热学说白了主要的内容还是热力学第一、第二定律那一套, 其他的讲的并不深入——甚至连自由能都没有深入讲解。

统计力学

统计力学的部分我是在 NUS 学的。这里教我们统计力学的是一位新老师, 她第一年带专业课。这位老师没有 ppt, 上课的时候直接把她写的讲义投到白板上, 用笔在旁边边写边讲。她的讲义里推导的地方基本都是留白, 上课讲解的时候她会边讲边在白板上补上空白的地方。这估计是为了上课教学而设计的。

虽然我感觉这位老师的讲义写得实在是很啰嗦, 但她真的是一位好老师。可能是每一位新老师的共同特征吧, 她很尊重学生的想法, 注重和学生的互动, 非常友善。这和国内的老师相比是非常不同的。

但是我依然没有听她讲课, 只是会在课下和她讨论问题聊聊天, 顺便也算是练习英语了。统计力学的部分我还是参看了 David Tong 的讲义居多, 然后在 2025 年 11 月写成了我这一份的讲义的前三章。

我的这份讲义包含了统计力学最基本的原理部分, 而没有涉及很多的应用部分。写这份讲义我有一个心得体会, 那就是理论物理学真的是为了给自然界找到一个底层的逻辑, 然后进行“演绎推理”; 但同时我也体会到, 自然界中的现象能够进行“优美地”演绎推理出来的其实是很小的一部分。有很多时候物理学家面对的其实是理想化的模型。其实上大学之后, 学了物理, 我体会到了对于好多的东西的讨论, 本来我觉得会很复杂但其实非常精炼 (比如能均分定理); 而对于某些东西的讨论我以为会很简单, 其实是比较复杂、没有一个很简洁精炼的推导的 (比如说化学反应速率常数等等)。

C.4 解锁天赋技能

再往后, 就要分研究方向了, 这就像是要解锁、选择自己的职业天赋技能一样。我选的是粒子物理、宇宙学这个方向, 跟统计力学的关系肯定是不如凝聚态这个方向这么密切了。当年在新手村里想着以后要好好学一学高中结构化学背后的原理 (好多都是凝聚态物理的内容) 的梦想, 估计是难以实现了。不过我相信, 这当一个兴趣爱好去了解了解还是很不错的。