

THERMAL PHYSICS & STATISTICAL MECHANICS

Y. Ding
September 23, 2025

CONTENTS

I	Thermal Physics	6
1	Energy in Thermal Physics	7
1.1	Thermal Equilibrium	7
1.2	The Ideal Gas	7
1.3	Heat and Work	9
1.3.1	Isothermal Compression	10
1.3.2	Adiabatic Compression	10
1.4	Heat Capacity	11
1.4.1	Latent Heat	12
1.4.2	Enthalpy	12
2	The Second Law	13
2.1	Two-State Systems	13
2.2	The Einstein Model of a Solid	14
2.3	Interacting Systems	15
2.4	Large Systems	15
2.4.1	Very Large Numbers	16
2.4.2	Stirling's Approximation	16
2.4.3	Multiplicity of a Large Einstein Solid	16
2.4.4	Sharpness of the Multiplicity Function	17
2.5	The Ideal Gas	18
2.5.1	Multiplicity of a Monatomic Ideal Gas	18
2.5.2	Interacting Ideal Gases	20
2.6	Entropy	20
2.6.1	Entropy of an Ideal Gas	21
2.6.2	Entropy of Mixing	21
3	Interactions and Implications	23
3.1	Temperature	23
3.2	Entropy and Heat	24
3.2.1	Predicting Heat Capacities	24
3.2.2	Measuring Entropies	25
3.3	Paramagnetism	25

3.3.1	Notation and Microscopic Physics	25
3.3.2	Numerical Solution	25
3.3.3	Analytic Solution	26
3.4	Mechanical Equilibrium and Pressure	27
3.4.1	The Thermodynamic Identity	28
3.5	Diffusive Equilibrium and Chemical Potential	28
4	Engines and Refrigerators	30
4.1	Heat Engines	30
4.1.1	Efficiency	30
4.1.2	Carnot Cycle	30
4.1.3	Another Example: Otto Cycle	31
4.2	Refrigerators	32
4.2.1	Definition	32
4.2.2	Second Law Revisited	32
4.3	Decade Thermometer	33
4.3.1	Isentropic Expansion	33
4.3.2	The Throttling Process	33
4.3.3	Liquefaction of Gases	34
4.3.4	Pumping	34
4.3.5	Dilution Refrigerator	34
4.3.6	Adiabatic Demagnetization	34
4.3.7	Laser Cooling	35
5	Free Energy	37
5.1	Free Energy as Available Work	37
5.1.1	Thermodynamic Potentials	37
5.1.2	Thermodynamic Identities	38
5.1.3	Second Derivatives and Maxwell Relations	39
5.2	Free Energy as a Force toward Equilibrium	39
5.2.1	Extensive and Intensive Quantities	40
5.2.2	Gibbs Free Energy and Chemical Potential	40
5.3	Phase Transformations of Pure Substances	41
5.3.1	Diamonds and Graphite	41
5.3.2	The Clausius-Clapeyron Relation	42
5.3.3	The van der Waals Model	42
5.4	Chemical Equilibrium	44

II	Statistical Mechanics	46
6	Boltzmann Statistics	47
6.1	The Boltzmann Factor	47
6.1.1	Boltzmann Distribution	47
6.1.2	The Partition Function	48
6.2	Average Values	48
6.2.1	Paramagnetism	49
6.2.2	Rotation and Vibration of Diatomic Molecules	49
6.3	The Equipartition Theorem	51
6.4	The Maxwell Speed Distribution	51
6.4.1	Speed Distribution	51
6.4.2	Energy Distribution	53
6.5	Partition Functions and Free Energy	53
6.6	Partition Functions for Composite Systems	54
6.7	Ideal Gas Revisited	54
6.7.1	The Partition Function	55
6.7.2	Predictions	56
7	Quantum Statistics	57
7.1	The Gibbs Factor	57
7.1.1	Average Values	57
7.1.2	The Grand Free Energy	58
7.2	Bosons and Fermions	58
7.2.1	The Distribution Functions	59
7.3	Degenerate Fermi Gas	60
7.3.1	Zero Temperature	60
7.3.2	Small Nonzero Temperatures	62
7.3.3	The Density of States	62
7.3.4	The Sommerfeld Expansion	63
7.4	Blackbody Radiation	63
7.4.1	The Planck Distribution	63
7.4.2	Total Particle Number and Energy	64
7.4.3	The Cosmic Background Radiation	66
7.4.4	Stefan-Boltzmann Law	66
7.5	Debye Theory of Solids	68
7.6	Bose-Einstein Condensation	69

7.6.1	The Condensation Temperature	70
7.6.2	Why Does it Happen?	72
8	Interacting Particles	74
8.1	Weakly Interacting Gases	74
8.1.1	The Partition Function	74
8.1.2	The Cluster Expansion	75
8.1.3	The Second Virial Coefficient	76
8.2	The Ising Model of a Ferromagnet	77
8.2.1	Exact Solution in One Dimension	78
8.2.2	The Mean Field Approximation	79
A	Proofs and Derivations	82
A.1	Energy and Entropy	82
A.1.1	Negative Heat Capacities	82
A.1.2	Multiplicity of a Low-Temperature Einstein Solid	83
A.1.3	Entropy of a Black Hole	83
A.1.4	Entropy of an Ising Paramagnet	84
A.2	Boltzmann Statistics	84
A.2.1	Energy of Quantum Harmonic Oscillators	84
A.2.2	Linear Degrees of Freedom	85
A.3	Quantum Statistics	85
A.3.1	Sommerfeld Expansion	85
A.3.2	Photon Gas Helmholtz Free Energy	87
A.3.3	Photon Spectrum in Terms of Wavelength	87
A.3.4	Below the Condensation Temperature	88
A.3.5	Bose Gas in a Harmonic Oscillator Potential	89

Thermodynamics and statistical mechanics deal with assemblies of large number ($\sim 10^{23}$) of atoms. It is nearly impossible to count and model this quantity of matter, let alone tracking their positions or velocities according to initial states. Thus, in thermal physics, we will abandon exact solutions. Instead, we focus on making approximations and dealing with statistical properties of matter (how they behave on average).

Part I

Thermal Physics

1 ENERGY IN THERMAL PHYSICS

1.1 Thermal Equilibrium

Temperature is an important concept in thermodynamics. There is an **operational definition** of temperature:

Temperature is what you measure with a thermometer.

For example, you want to know the temperature of a soup. You put the mercury thermometer into the soup, and after a while, the thermometer and the soup will end up with the same temperature. The reading on the thermometer (based on the volume expansion of mercury) tells the temperature. This brings the **theoretical definition** of temperature:

Temperature is the thing that is the same for two objects, after they have been in contact long enough.

After two objects have been in contact long enough, they are in **thermal equilibrium**. The time required for a system to come to thermal equilibrium is called the **relaxation time**. By “contact” we mean the two objects can exchange energy spontaneously. There are other types of equilibrium which corresponds to different exchange quantity and measurements.

Exchange quantity	Type of equilibrium	measurements
energy	thermal	temperature
volume	mechanical	pressure
particles	diffusive	chemical potential

Now, when two objects can exchange energy, the object that releases energy is at higher temperature and the object that receives energy is at lower temperature. Here comes the more precise theoretical definition of temperature:

Temperature is a measure of the tendency of an object to spontaneously give up energy to its surroundings. When two objects are in thermal contact, the one that tends to spontaneously lose energy is at the higher temperature.

The temperature has a minimum value called the **absolute zero**, which is about -273°C , or 0 K. The pressure of any gas will approach zero as the temperature approaches absolute zero. In most of the equations of thermodynamics, the scale is in the **Kelvin scale**, or the absolute temperature scale.

1.2 The Ideal Gas

The ideal gas has the following property:

- Ideal gases usually have high temperature and low pressure. There are no intermolecular forces so that the molecules do not attract each other.
- Molecules are point-like and have zero size.
- The only interactions between molecules are elastic collisions.

For a low-density gas, many of its properties are summarized in the **ideal gas law**:

$$PV = NkT, \tag{1.1}$$

where P is the pressure, V is the volume, N is the number of particles in the gas, and $k \approx 1.381 \times 10^{-23}$ J/K is the **Boltzmann constant**.

Using the ideal gas law, we can relate the concept of temperature to the exchange of energy between two systems. Let us now consider a microscopic model of a gas with only one gas molecule in a box shown in Figure 1.1.

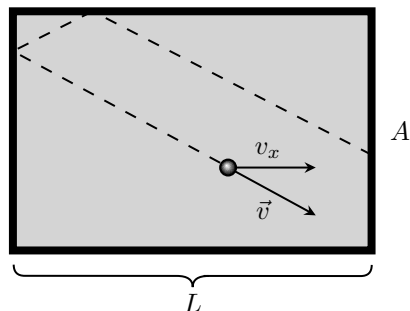


Figure 1.1: One gas molecule in the box.

Assume the cross sectional area of the box is A and its length is L . The molecule has a velocity \mathbf{v} with horizontal component v_x . The collisions are elastic so no kinetic energy is lost and the speed of the molecule is constant. Now, the pressure exerted on the wall by the molecule at most of the time is zero, but it collides with the wall for a brief amount of time periodically. We want to know the “average pressure” on the right side wall over a long time period:

$$\bar{P} = \frac{\bar{F}_{x,\text{on piston}}}{A} = \frac{-\bar{F}_{x,\text{on molecule}}}{A} = -\frac{m}{A} \left(\frac{\Delta v_x}{\Delta t} \right),$$

by Newton’s third law and second law. Let us take Δt to be the time for the molecule to complete a round-trip (from left to right to left), so

$$\Delta t = 2L/v_x.$$

The velocity change in x -direction is

$$\Delta v_x = (-v_x) - v_x = -2v_x.$$

Hence the average pressure on the piston is

$$\bar{P} = \frac{m}{A} \frac{(-2v_x)}{2L/v_x} = \frac{mv_x^2}{AL} = \frac{mv_x^2}{V} \implies \bar{P}V = mv_x^2.$$

If there are N molecules in the box, the pressure is rather continuous so we drop the overbar on the P . Moreover,

$$PV = mv_{1x}^2 + mv_{2x}^2 + mv_{3x}^2 + \cdots = Nm\overline{v_x^2},$$

where $\overline{v_x^2}$ is the average of the molecules’ v_x^2 . Using the ideal gas law (1.1),

$$kT = m\overline{v_x^2} \quad \text{or} \quad \overline{\frac{1}{2}mv_x^2} = \frac{1}{2}kT.$$

There is nothing special about x , this equation is also true in y and z -directions, so

$$\overline{\frac{1}{2}mv_y^2} = \overline{\frac{1}{2}mv_z^2} = \frac{1}{2}kT.$$

Then we obtain the average translational kinetic energy of the molecules:

$$\bar{K}_{\text{trans}} = \overline{\frac{1}{2}mv^2} = \frac{1}{2}m(\overline{v_x^2} + \overline{v_y^2} + \overline{v_z^2}) = \frac{3}{2}kT.$$

This equation is a special case of the **equipartition theorem**, which states that: at temperature T , the average energy of any quadratic degree of freedom is $\frac{1}{2}kT$. By quadratic degrees of freedom we mean the quadratic function of a coordinate or velocity component. Many kinds of energy follow the quadratic function form, such as translational kinetic energy $\frac{1}{2}mv_x^2$, rotational kinetic energy $\frac{1}{2}I\omega_x^2$, or elastic potential energy $\frac{1}{2}kx^2$. If a system contains N molecules, each with f degrees of freedom, and there are no other (non-quadratic) temperature-dependent forms of energy, then its total thermal energy is

$$U_{\text{thermal}} = N \cdot f \cdot \frac{1}{2}kT.$$

Example 1.1. Diatomic Molecule

The diatomic molecules (e.g. O_2) have different average energy from monatomic ones. Particularly, the average energy of diatomic molecules is

$$U = \frac{7}{2}kT, \quad [f = 7].$$

Like monatomic molecules, there are 3 degrees of freedom related to translational kinetic energy $\frac{1}{2}mv_i^2$, $i = x, y, z$. Suppose the molecule is aligned with the x -axis. It has 2 more degrees of freedom from rotational kinetic energy about the y - and z -axis, $\frac{1}{2}I\omega_y^2$ and $\frac{1}{2}I\omega_z^2$. Moreover, the two atoms in the molecule are vibrating, giving rise to spring potential energy kx^2 and kinetic energy of the two atoms. Hence in total there are 7 quadratic degrees of freedom possible. However, to excite the rotational and vibrational quadratic degrees of freedom, the gas needs to have a very high temperature.

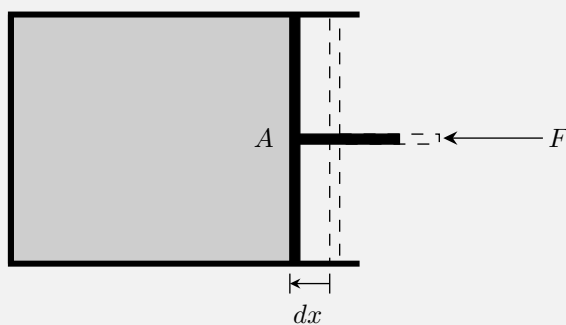
1.3 Heat and Work

Heat is a flow of energy from two systems which is caused by a difference in temperature. Heat transfer can be due to conduction, convection, and radiation. **Work** is any other transfer of energy (i.e. not caused by a difference in temperature). The **first law of thermodynamics** says that the total change in internal energy U of a system is the heat Q transferred into, and the work W done on the system:

$$\Delta U = Q + W, \quad (1.2)$$

Example 1.2. Compression

Consider a box of ideal gas with a piston of area A pushing with a force F .



The work element done on the gas by displacing the piston to the left by dx is just

$$dW = F dx,$$

In terms of quantities useful in gas, we have the pressure and volume relation

$$dW = PA dx = -P dV.$$

The total work done on the gas is the integral of the work element:

$$W = - \int_{V_i}^{V_f} P dV.$$

In this compression process we shall assume that the piston is moving reasonably slow so that the pressure is uniform inside (which is called **quasistatic**).

The compression of an ideal gas has two important cases:

- Isothermal: if we compress slowly, waiting the heat to exchange with the environment, the temperature remains constant.
- Adiabatic: if we compress quickly, or the gas is isolated from the environment so that there is no heat flow, the temperature of the gas increases.

1.3.1 Isothermal Compression

By the ideal gas law (1.1),

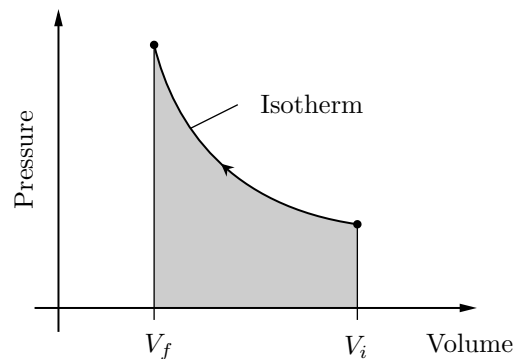
$$P = NkT/V,$$

The work done on the gas is

$$W = - \int_{V_i}^{V_f} P dV = -NkT \int_{V_i}^{V_f} \frac{1}{V} dV = NkT \ln \left(\frac{V_i}{V_f} \right). \quad (1.3)$$

In this case, by the first law of thermodynamics (1.2) and the fact that the internal energy is proportional to temperature ($\Delta U = 0$ isothermally), the heat transfer into the system is

$$Q = \Delta U - W = -W = NkT \ln \left(\frac{V_f}{V_i} \right).$$



1.3.2 Adiabatic Compression

In an adiabatic compression, there is no heat flowing in or out of the system,

$$Q = 0.$$

Hence by the first law of thermodynamics (1.2) the change in internal energy is just the work done on the gas,

$$\Delta U = W.$$

The adiabatic curve moves from one isotherm to another. The infinitesimal change in internal energy is

$$dU = \frac{1}{2} f Nk dT = dW = -P dV.$$

Plugging in $P = NkT/V$, it is a differential equation relating temperature and volume:

$$\frac{f}{2} \frac{dT}{T} = -\frac{dV}{V}.$$

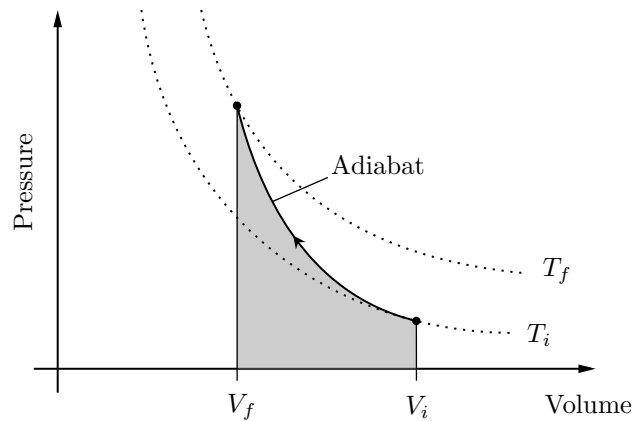
We will obtain the following result depending on the number of quadratic degrees of freedom,

$$V_f T_f^{f/2} = V_i T_i^{f/2} \quad \text{or} \quad VT^{f/2} = \text{const.}$$

If it is to be expressed in terms of pressure, substitute $T = PV/Nk$ for an ideal gas so that

$$V_f^{1+f/2} P_f^{f/2} = V_i^{1+f/2} P_i^{f/2} \quad \text{or} \quad V^\gamma P = \text{const.}$$

where $\gamma = (f + 2)/f$.



In an adiabatic curve, the initial point moves to the final point from two isotherms T_i and T_f .

1.4 Heat Capacity

How much heat is needed to raise the temperature of an object by ΔT ? The answer is the equation

$$Q = C\Delta T \quad (1.4)$$

where C is called the **heat capacity**. This can be rewritten as

$$C = \frac{Q}{\Delta T}.$$

- If there is no work, $W = 0$, then the volume of the gas is constant. We will get the heat capacity at constant volume:

$$C_V = \left. \frac{\Delta U}{\Delta T} \right|_V = \left. \frac{\partial U}{\partial T} \right|_V.$$

Suppose we are considering the heat capacity of an ideal gas with quadratic degrees of freedom f ,

$$C_V = \frac{\partial U}{\partial T} = \frac{\partial}{\partial T} \left(\frac{1}{2} N f k T \right) = \frac{N f k}{2}.$$

For example, the heat capacity of He or Ar is $\frac{3}{2} N k \approx 12.5 \text{ J}/(\text{K} \cdot \text{mol})$ with $f = 3$. Diatomic molecules like oxygen gas O_2 will have higher heat capacity due to more degrees of freedom, and polyatomic molecules will have even more. Their available degrees of freedom also depends on the temperature of the gas according to Figure 1.2.

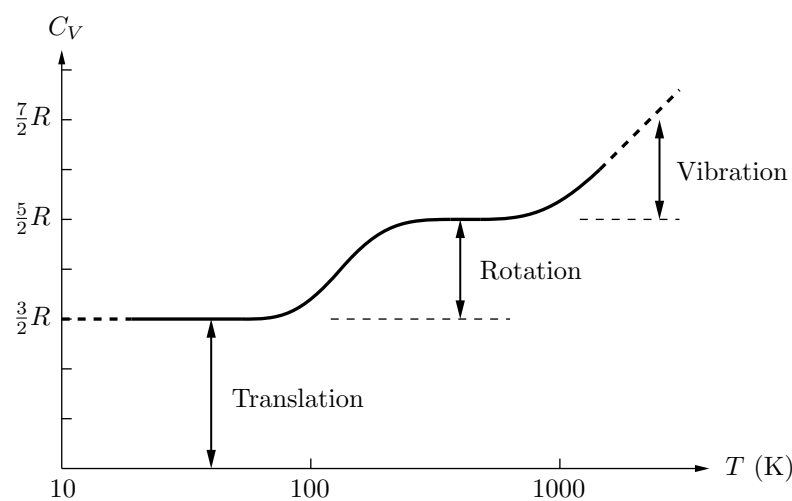


Figure 1.2: Heat capacity at constant volume of one mole of hydrogen (H_2) gas ($R = 6.02 \times 10^{23} k$ is the ideal gas constant). Below about 100 K only the three translational degrees of freedom are active. Around room temperature the two rotational degrees of freedom are active. Above 1000 K the two vibrational degrees of freedom also become active.

- If the pressure is constant, the heat capacity is

$$C_P = \left. \frac{Q}{\Delta T} \right|_P = \left. \frac{\Delta U - (-P\Delta V)}{\Delta T} \right|_P = \left. \left(\frac{\partial U}{\partial T} + P \frac{\partial V}{\partial T} \right) \right|_P.$$

Again, for an ideal gas with degrees of freedom f ,

$$C_P = \frac{\partial U}{\partial T} + P \frac{\partial V}{\partial T} = \frac{Nfk}{2} + P \frac{\partial}{\partial T} \left(\frac{NkT}{P} \right) = C_V + Nk,$$

where in the second equality we use the ideal gas law (1.1). Note that the heat capacity at constant pressure differs from heat capacity at constant volume only by a constant Nk that does not depend on the pressure.

What about solids? A well behaved solids can be approximated as a lattice with individual atoms in 3D simple harmonic oscillator. Hence its quadratic degrees of freedom is 6 (3 kinetic and 3 potential). The heat capacity per mole is

$$\frac{6}{2}R = 3R \approx 25 \text{ J/(K} \cdot \text{mol)}.$$

Such result is called the **rule of Dulong and Petit** and it works at high temperatures. The word “high” varies among different materials: the room temperature is considered “high” for metals like copper, aluminum, or lead, but “low” for diamond. Heat capacities are normally positive, but there is an important class of exceptions: systems of particles held together by gravity, such as stars and star clusters (see Appendix A.1.1).

1.4.1 Latent Heat

During phase transformation, like from liquids to gases, the temperature does not change but you are still adding heat to the system. In this case, the heat capacity seems to be infinite, $C = Q/\Delta T = \infty$. We need another quantity called the **specific latent heat**, which is similar to the heat capacity to describe how much heat is needed, but for a phase change,

$$L \equiv \frac{Q}{M} \tag{1.5}$$

If you know the mass M of the object undergoing the phase transformation, you can calculate the heat required.

1.4.2 Enthalpy

In the natural world and in the lab, we often deal with processes at constant pressures (like 1 atm). Keeping track of the compression-expansion work done during these processes might be tedious. It is convenient to define the **enthalpy** H as

$$H \equiv U + PV. \tag{1.6}$$

In some sense, the enthalpy is the energy of the system U plus the energy needed to *make a room* for it, PV , at a constant pressure. In other words, the enthalpy gives you the energy to create a system out of nothing. For some constant pressure process,

$$\Delta H = \Delta U + P\Delta V.$$

By the first law of thermodynamics (1.2),

$$\Delta U = Q + (-P\Delta V) + W_{\text{other}},$$

which says that $\Delta H = Q + W_{\text{other}}$, where W_{other} is the work not done by compression or expansion of a gas. If $W_{\text{other}} = 0$, $\Delta H = Q$, then the heat capacity, or “enthalpy capacity” is

$$C_P = \left. \frac{\partial H}{\partial T} \right|_P.$$

The enthalpy so useful in chemistry that there are tables of enthalpy of chemical reactions.

2 THE SECOND LAW

This chapter deals with the second law of thermodynamics and the spontaneity of thermodynamic processes. The second law has a big question in it: why does heat flow spontaneously from a hotter object to a cooler object and reach an equilibrium, but never around. An even bigger question is why many thermodynamic processes happen in one direction but never the reverse. Actually, the word “never” should be changed to “very unlikely”, but what does “very” mean in thermodynamics? The answer is

$$\text{very unlikely} \sim 1/10^{23}!$$

so we can basically rule out the probability of many reverse processes. To understand why this is the case, we need to study how systems store energy and to count how the energy can be arranged using [combinatorics](#).

2.1 Two-State Systems

A [microstate](#) is a specific state of system among all possible states. A [macrostate](#) is a more general state which contains (or which is a set of) microstates. If you know the microstate of the system, then you also know its macrostate, but not vice versa. The number of microstates corresponding to a given macrostate is called the [multiplicity](#) of that macrostate, Ω .

Example 2.1. Flipping Coins

Suppose there are three coins (a penny, a nickel, and a dime) to flip and we label the outcomes according to heads and tails. There are totally 8 equally possible outcomes as listed in Table 2.1. We say that there are 8 microstates as the total outcome. There are, for example, 3 different ways to get exactly two heads and a tail (HHT). The macrostate in this case is [two heads] and there are 3 microstates corresponding to such macrostate. Hence the multiplicity of this macrostate is 3, or $\Omega(2 \text{ heads}) = 3$. We can count the multiplicity of other macrostates to be: $\Omega(3 \text{ heads}) = \Omega(0 \text{ heads}) = 1$, $\Omega(1 \text{ head}) = 3$. The total multiplicity of all four macrostates is 8, the total number of microstates, $\Omega(\text{all}) = 8$. The probability of getting any particular macrostate is then

$$P(n) = \frac{\Omega(n)}{\Omega(\text{all})}.$$

Penny	Nickel	Dime
H	H	H
H	H	T
H	T	H
T	H	H
H	T	T
T	H	T
T	T	H
T	T	T

Table 2.1: A list of all possible microstates of a set of three coins.

Now instead of 3 coins, we have 100 coins to flip. For the 0-heads macrostate, its multiplicity is $\Omega(0) = 1$, i.e. there is only one outcome that no heads appear. If there were exactly one head, the multiplicity is $\Omega(1) = 100$. If there were exactly two heads, we can see this as choosing two coins to turn heads-up from 100 coins. We first have 100 choices for the first coin, then 99 for the second. We can choose the pair in any order (that is, choosing coin A then B is the same as choosing B then A) so we divide the total choices by 2:

$$\Omega(2) = \frac{100 \cdot 99}{2}.$$

Similarly,

$$\Omega(3) = \frac{100 \cdot 99 \cdot 98}{3 \cdot 2}.$$

By this pattern,

$$\Omega(n) = \frac{100 \cdot 99 \cdots (100 - n + 1)}{n \cdots 2 \cdot 1} = \frac{100!}{n!(100 - n)!} \equiv \binom{100}{n}.$$

Generalize this to the case where there are N total coins instead of 100. The multiplicity of the macrostate with n heads is

$$\Omega(N, n) = \frac{N!}{n!(N-n)!} = \binom{N}{n}.$$

2.2 The Einstein Model of a Solid

Now move to a more physical system, a collection of microscopic systems that can each store any number of energy “units”. Such microscopic system can be a quantum harmonic oscillator with potential energy $\frac{1}{2}k_s x^2$. Quantum harmonic oscillators have energy level in equal spaces, or same energy units $\hbar\omega = hf$, where h is the Planck’s constant and f is the natural frequency of the oscillator $f = \frac{1}{2\pi}\sqrt{k_s/m}$ (see Figure 2.1).

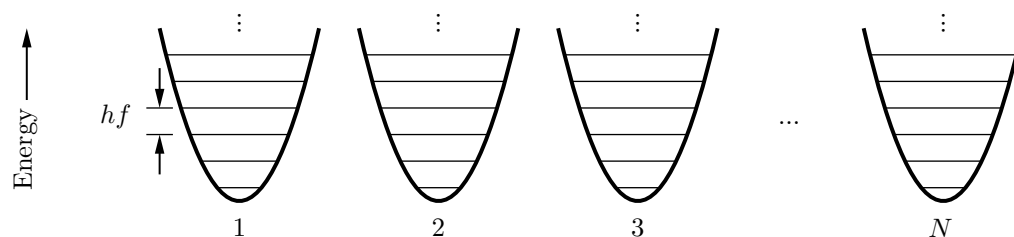


Figure 2.1: In quantum mechanics, any system with a quadratic potential energy function has evenly spaced energy levels separated in energy by hf . An Einstein solid is a collection of N such oscillators, all with the same frequency.

Harmonic oscillators are common in solids. In a three-dimensional solid, each atom can oscillate in three independent directions. Hence if there are N oscillators, there are $N/3$ atoms. The model of a solid composed of identical oscillators is called an **Einstein solid**.

Suppose there are $N = 3$ harmonic oscillators in an Einstein solid. We can write out the energy of each oscillator (in units of $\hbar\omega$) for each solid and study their combinations as microstates and macrostates. This time we consider the total energy to be macrostates, according the following table.

Oscillator:	#1	#2	#3	Oscillator:	#1	#2	#3
$E = 0$	0	0	0	$E = 3$	3	0	0
$E = 1$	1	0	0		0	3	0
	0	1	0		0	0	3
	0	0	1		2	1	0
$E = 2$	2	0	0		2	0	1
	0	2	0		1	2	0
	0	0	2		0	2	1
	1	1	0		1	0	2
	1	0	1		0	1	2
	0	1	1		1	1	1

For example, there is 1 microstates for $E = 0$, which is zero for all oscillators:

$$\Omega(0) = 1.$$

There are 3 microstates that have harmonic oscillators sharing the energy $E = 1$, $\Omega(1) = 3$, etc. The general formula for the multiplicity of an Einstein solid with N oscillators and q energy units is

$$\Omega(N, q) = \binom{q + N - 1}{q} = \frac{(q + N - 1)!}{q!(N - 1)!}. \quad (2.1)$$

2.3 Interacting Systems

Consider two Einstein solids that are in thermal contact and are isolated from other systems. There are N_A oscillators in solid A with energy q_A and N_B oscillators in solid B with energy q_B . The total energy in two solids is

$$q_{\text{total}} = q_A + q_B.$$

Example 2.2. A very small system

A very small system could be two solids in thermal contact, each with three harmonic oscillators and they contain a total of six energy units:

$$N_A = N_B = 3, \quad q_{\text{total}} = q_A + q_B = 6.$$

There are 7 macrostates, $q_A = 0, 1, 2, \dots, 6$ (once we know q_A , we will know q_B because the total energy is fixed). Now count or compute the multiplicities using $\Omega(N, q) = \binom{q+N-1}{q}$. The result is listed in Figure 2.2.

q_A	Ω_A	q_B	Ω_B	$\Omega_{\text{total}} = \Omega_A \Omega_B$
0	1	6	28	28
1	3	5	21	63
2	6	4	15	90
3	10	3	10	100
4	15	2	6	90
5	21	1	3	63
6	28	0	1	28
				total = 462

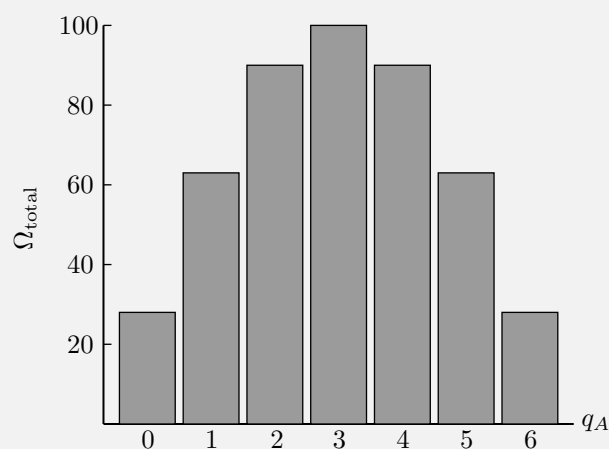


Figure 2.2: Macrostates and multiplicities of a system of two Einstein solids, each containing three oscillators, sharing a total of six energy units.

Let us now make a big assumption: All states are equally likely. In thermodynamics, it means

For an isolated system in thermal equilibrium, all microstates are equally probable.

This is called the **fundamental assumption of statistical mechanics**. It is valid for a **microcanonical ensemble**, a system where the total energy, particle number, and volume are specified. Here is the point: though all microstates are equally likely, some *macrostate* are more likely than others because of the multiplicity. In the previous example, the behavior is already shown. The probability of getting equal energy in A and B is higher than the probability of getting all energies in B . To see how this relates to the irreversibility, let us have some bigger numbers.

2.4 Large Systems

You may think of some “large” systems with $N, q \approx$ few hundred or $N, q \approx$ few thousand and check their multiplicity. The multiplicity graph of such “large” systems shows some trends in Figure 2.3. However, solids in daily life contains $\sim 10^{23}$ oscillators, and this graph is too sharp to draw. We will show that at this level, only a tiny fraction of all macrostates are reasonably possible. Thermodynamically, suppose we start the system with all energy in B and watch the system evolve. If the system is rearranging itself randomly, we will see that the energy will flow from B to A almost certainly and it will not flow back. Before we do some formulation, we need to study the mathematics of very large numbers.

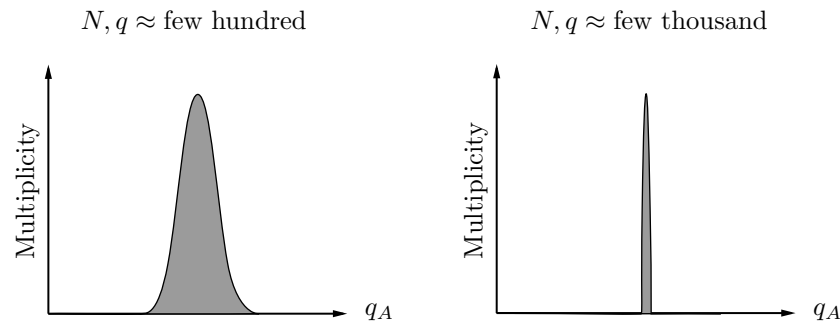


Figure 2.3: Typical multiplicity graphs for two interacting Einstein solids, containing a few hundred oscillators and energy units (left) and a few thousand (right).

2.4.1 Very Large Numbers

There are three kinds of numbers in statistical mechanics: small numbers, large numbers, and very large numbers.

- Small numbers are normal numbers like 2, e , or 23.
- Large numbers are usually made by exponentiating small numbers. The most important one is the Avogadro's number, $\sim 10^{23}$. The property of a large number is that we can add a small number to a large number without changing it,

$$10^{23} + 23 = 10^{23}.$$

- Very large numbers are often made by exponentiating large numbers, such as $10^{10^{23}}$. You can multiply a very large number by a large number without changing it,

$$10^{10^{23}} \times 10^{23} = 10^{(10^{23}+23)} = 10^{10^{23}}.$$

2.4.2 Stirling's Approximation

When calculating multiplicities, we often encounter factorials. To evaluate factorials of large numbers, we can use [Stirling's approximation](#):

$$n! \approx n^n e^{-n} \sqrt{2\pi n}, \quad (2.2)$$

for $n \gg 1$. Sometimes Stirling's approximation is written as

$$\ln n! = n \ln n - n. \quad (2.3)$$

Here is a simple proof of the Stirling's approximation. Take the logarithm of $n!$:

$$\ln n! = \ln[n \cdot (n-1) \cdot (n-2) \cdots 1] = \ln n + \ln(n-1) + \ln(n-2) + \cdots + \ln 1.$$

For large n , this is just like an integral of $\ln x$,

$$\ln n! \approx \int_0^n \ln x \, dx = n \ln n - n \implies n! \approx n^n e^{-n}.$$

The $\sqrt{2\pi n}$ factor can be omitted if we are considering large numbers. A more careful proof needs to make use of the Gamma function

$$n! = \Gamma(n+1) = \int_0^\infty x^n e^{-x} dx.$$

2.4.3 Multiplicity of a Large Einstein Solid

Consider an Einstein solid with N oscillators with q energy units. Suppose N is large, and $q \gg N$ (high temperature limit). The multiplicity of such a system is often a very large number because it contains factorials of a large number (such as $N!$). Using the

property of very large numbers, the multiplicity of each macrostate with energy q is given by (2.1),

$$\Omega(N, q) = \binom{q + N - 1}{q} = \frac{(q + N - 1)!}{q!(N - 1)!} \approx \frac{(q + N)!}{q!N!}.$$

Take the logarithm and use Stirling's approximation (2.3):

$$\begin{aligned} \ln \Omega &= \ln \left[\frac{(q + N)!}{q!N!} \right] = \ln(q + N)! - \ln q! - \ln N! \\ &\approx (q + N) \ln(q + N) - (q + N) - q \ln q + q - N \ln N + N \\ &= (q + N) \ln(q + N) - q \ln q - N \ln N. \end{aligned}$$

The first logarithm can be simplified by the high temperature limit,

$$\ln(q + N) = \ln \left[q \left(1 + \frac{N}{q} \right) \right] = \ln q + \ln \left(1 + \frac{N}{q} \right) \approx \ln q + \frac{N}{q},$$

where we use the Taylor expansion of $\ln(1 + x) \approx x$ for $|x| \ll 1$. Hence

$$\ln \Omega \approx N \ln \frac{q}{N} + N + N \left(\frac{N}{q} \right).$$

The last term is negligible compared to other terms. We will finally find that

$$\Omega(N, q) \approx e^{N \ln(q/N)} e^N = \left(\frac{eq}{N} \right)^N. \quad (2.4)$$

In this formula, not only the term in the bracket is a large number, the exponent is also a large number, so the multiplicity is a very large number. For multiplicity of a low temperature Einstein solid ($q \ll N$), see Appendix A.1.2.

2.4.4 Sharpness of the Multiplicity Function

The multiplicity of a combined system A and B with total energy q , each with N oscillators, is

$$\Omega = \left(\frac{eq_A}{N} \right)^N \left(\frac{eq_B}{N} \right)^N = \left(\frac{e}{N} \right)^{2N} (q_A q_B)^N. \quad (2.5)$$

where q_A and q_B are number of energy units in solids A and B respectively. The graph of multiplicity is apparently a sharp spike, and the spike is at location $q_A = q/2$ where energy is evenly distributed between the two solids. The height of this peak is then

$$\Omega_{\max} = \left(\frac{e}{N} \right)^{2N} \left(\frac{q}{2} \right)^{2N}.$$

Near the peak, say $q_A = q/2 + x$ and $q_B = q/2 - x$, the multiplicity according to (2.5) is

$$\Omega = \left(\frac{e}{N} \right)^{2N} \left[\left(\frac{q}{2} \right)^2 - x^2 \right]^N.$$

It is again useful to take the logarithm and use Taylor expansion for the second factor

$$\begin{aligned} \ln \left[\left(\frac{q}{2} \right)^2 - x^2 \right]^N &= N \ln \left[\left(\frac{q}{2} \right)^2 - x^2 \right] = N \ln \left[\left(\frac{q}{2} \right)^2 \left(1 - \left(\frac{2x}{q} \right)^2 \right) \right] \\ &= N \left[\ln \left(\frac{q}{2} \right)^2 + \ln \left(1 - \left(\frac{2x}{q} \right)^2 \right) \right] \\ &\approx N \left[\ln \left(\frac{q}{2} \right)^2 - \left(\frac{2x}{q} \right)^2 \right]. \end{aligned}$$

The multiplicity near the peak turns out to be a Gaussian distribution:

$$\Omega = \Omega_{\max} e^{-N(2x/q)^2} \quad \text{with} \quad \sigma = \frac{q}{2\sqrt{N}},$$

where

$$\Omega_{\max} = \left(\frac{e}{N}\right)^{2N} e^{N \ln(q/2)} = \left(\frac{e}{N}\right)^{2N} \left(\frac{q}{2}\right)^{2N}.$$

Let us assume $N = 10^{20}$, one standard deviation (the width) is only ten billionth of the entire scale of the graph. But at only 10σ , the multiplicity is already a factor of $e^{-100} \approx 10^{-44}$ smaller than the maximum value.

Here comes the conclusion: for thermodynamically large systems at equilibrium, any random fluctuations away from the most likely macrostate will be utterly unmeasurable. That is, the fluctuations away from the most likely macrostate never occur, which is called the **thermodynamic limit**. At this limit, second law becomes a *law*, not a probabilistic statement.

2.5 The Ideal Gas

The conclusion from the previous section turns out to be true for any interacting objects, with the assumption of large number of particles and energy units. In this section, we will see how this is true for an ideal gas.

2.5.1 Multiplicity of a Monatomic Ideal Gas

Consider a monatomic noble ideal gas (Ar, He, etc). What is the multiplicity Ω of one gas atom with mass m and with kinetic energy U in a volume V ? Classically, the answer looks like infinity because the atom can be anywhere and the velocity can point in any direction. We shall apply quantum mechanics to solve this problem.

First, let us work on this problem in the proportionality aspect. If the volume is expanded to $2V$, there should be twice as many states for the atom. Similarly, if the volume of the **momentum space** V_p doubles, it just means the atom can choose twice as many momentum vectors. (The momentum space is just a space with axes p_x , p_y , and p_z . Each point in the momentum space corresponds to a momentum vector for the particle.) Hence the multiplicity is proportional to both the volume V , and the volume V_p in the momentum space:

$$\Omega_1 \propto VV_p$$

where Ω_n is the multiplicity for n molecules (in this case atoms).

Now we determine the volume of the momentum space V_p . The momentum of the molecule is related to the kinetic energy by

$$p_x^2 + p_y^2 + p_z^2 = 2mU.$$

which is a sphere of radius $\sqrt{2mU}$ in the momentum space (see Figure 2.4). This means the momentum of the molecule is confined to this surface, i.e. the “volume” of momentum space is the surface area of this sphere (maybe times a small thickness so that it is a volume). To find the proportionality constant, we invoke the Uncertainty principle:

$$\Delta x \Delta p_x \gtrsim h.$$

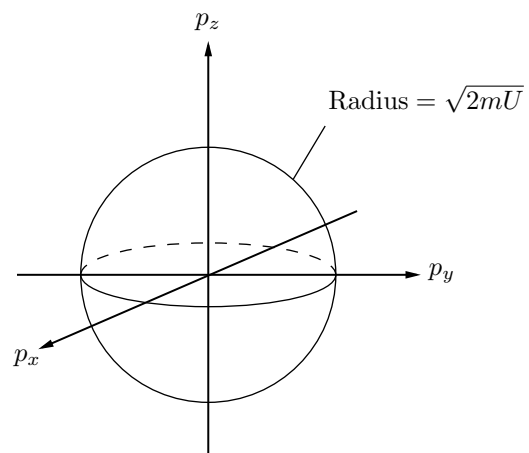


Figure 2.4: A sphere in momentum space with radius $\sqrt{2mU}$.

(Do not worry about whether it is h or $\hbar/2$. We are dealing with large numbers so these factors are unimportant for now. Also, we will only take the $\sim h$ part but not greater because we want to specify the position and momentum as precisely as possible.)

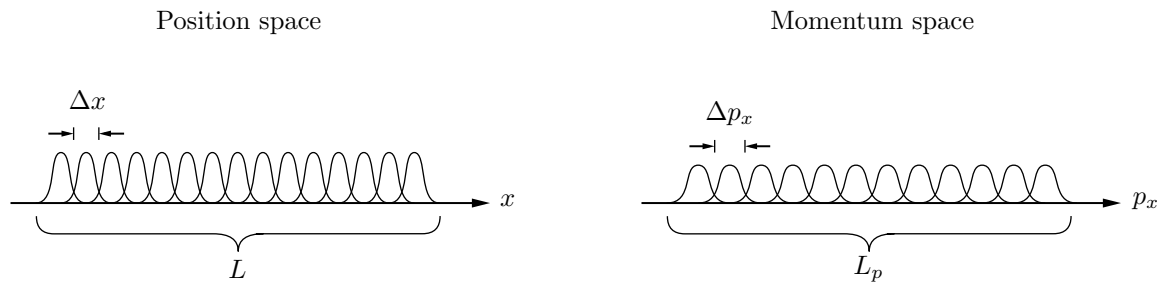


Figure 2.5: A number of “independent” position states and momentum states for a quantum-mechanical particle moving in one dimension.

The number of independent wavefunctions is finite if the total available position space and momentum space are limited. A one-dimensional example is shown in Figure 2.5. The numbers of distinct position states and momentum states are respectively $L/(\Delta x)$ and $L_p/(\Delta p_x)$. The number of distinct states is then

$$\frac{L}{\Delta x} \frac{L_p}{\Delta p_x} = \frac{LL_p}{h}.$$

The same works for y and z directions, so the multiplicity is

$$\Omega_1 = \frac{VV_p}{h^3}.$$

You may think of the two molecule case where there are Ω_1 states for molecule 1 and also Ω_1 states for molecule 2. Then the total multiplicity is just the product Ω_1^2 . It is not correct due to a more complicated momentum space V_p . Recall that only the *total* energy of the two molecules is constrained. Assume the molecules have the same mass,

$$p_{1x}^2 + p_{1y}^2 + p_{1z}^2 + p_{2x}^2 + p_{2y}^2 + p_{2z}^2 = 2mU.$$

This defines a surface of a “six-dimensional hypersphere” in the six-dimensional momentum space. Hence the total multiplicity is

$$\Omega_2 = \frac{V^2}{h^6} \times (\text{The area of a 6-dimensional hypersphere with radius } \sqrt{2mU}).$$

In most cases, particles are indistinguishable. That is, interchanging molecules with each other gives the same state, so we are overcounting the microstates twice. The multiplicity of a gas with two indistinguishable molecule is

$$\Omega_2 = \frac{1}{2} \frac{V^2}{h^6} \times (\text{The area of a 6-dimensional hypersphere with radius } \sqrt{2mU}).$$

Generalizing to N particles

$$\Omega_N = \frac{1}{N!} \frac{V^N}{h^{3N}} \times (\text{The area of a } 3N\text{-dimensional hypersphere with radius } \sqrt{2mU}). \quad (2.6)$$

To know what exactly the area of hyperspheres is, we need some formula. The area of a d -dimensional hypersphere with radius R is given by

$$A = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} R^{d-1} \quad (2.7)$$

where $\Gamma(n) = (n-1)!$ is the gamma function. For half-integer reference, $(1/2)! = \sqrt{\pi}/2$. Plug (2.7) into the multiplicity formula (2.6)

$$\Omega_N = \frac{1}{N!} \frac{V^N}{h^{3N}} \frac{2\pi^{3N/2}}{(\frac{3N}{2}-1)!} (\sqrt{2mU})^{3N-1} \approx \frac{1}{N!} \frac{V^N}{h^{3N}} \frac{\pi^{3N/2}}{(\frac{3N}{2})!} (\sqrt{2mU})^{3N}. \quad (2.8)$$

We can think of the multiplicity as

$$\Omega_N = f(N) V^N U^{3N/2}$$

which shows its relation with the volume and total kinetic energy with some function $f(N)$ that depends only on N . Ω_N is a very large number, so we can omit some large factors. You may think this proof is not a rigorous one, as we are seeing area as volume, interchanging h and \hbar , etc. Later, we will have a more rigorous derivation using Boltzmann Statistics.

2.5.2 Interacting Ideal Gases

Consider two ideal gases A and B which are thermally interacting each other but are separated by a partition (see Figure 2.6). Assume that they both have N molecules of the same species.

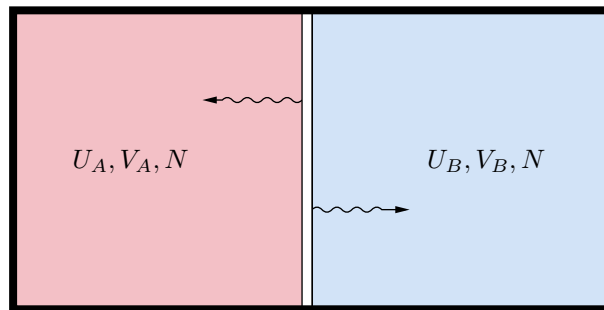


Figure 2.6: Two ideal gases, each confined to a fixed volume, separated by a partition that allows energy to pass through. The total energy of the two gases is fixed.

Like we did for Einstein solids, the total multiplicity is the product of multiplicities of individual gas,

$$\Omega_{\text{total}} = \Omega_A \Omega_B = [f(N)]^2 (V_A V_B)^N (U_A U_B)^{3N/2}.$$

Again, this multiplicity as a function of U_A is very sharply peaked. The same reasoning applies here: the peak is located at $U_A = U_{\text{tot}}/2$. We will arrive at a Gaussian distribution. The width of the peak (the standard deviation) is

$$\sigma = \frac{U_{\text{total}}}{\sqrt{3N/2}}.$$

We can also let the gases to exchange volume with each other by making the partition movable. The multiplicity as a function of V_A is very sharply peaked:

$$\sigma = \frac{V_{\text{total}}}{\sqrt{N}}.$$

2.6 Entropy

The second law of thermodynamics says that the multiplicity tends to increase. As we have observed, multiplicity are usually very large numbers. It is natural to take logarithms for convenience. We can define the **entropy** as the boltzmann constant times the logarithm of multiplicity.

$$S \equiv k \ln \Omega. \quad (2.9)$$

It has a unit of J/K. For example, the multiplicity of an einstein solid is $\Omega = (eq/N)^N$ at the high temperature limit ($q \gg N$) is Its entropy is just

$$S = k \ln \left(\frac{eq}{N} \right)^N = Nk \left[\ln \left(\frac{q}{N} \right) + 1 \right].$$

Let us say $N = 10^{22}$ and $q = 10^{24}$. The entropy is about 0.77 J/K, which is a small and reasonable number. Unlike multiplicity, entropy is additive

$$S_{\text{total}} = k \ln \Omega_{\text{total}} = k \ln (\Omega_A \Omega_B) = k \ln \Omega_A + k \ln \Omega_B = S_A + S_B.$$

Moreover, if the multiplicity is maximized, the entropy is maximized because $\ln x$ increase monotonically with x . Hence the **second law of thermodynamics** can be restated as:

Any large system in equilibrium will be found in the macrostate with the greatest entropy, aside from fluctuations that are normally too small to measure. Entropy tends to increase.

2.6.1 Entropy of an Ideal Gas

The entropy of a monatomic ideal gas (after some Stirling's approximation to Equation 2.8) is

$$S = Nk \left[\ln \left(\frac{V}{N} \left(\frac{4\pi m U}{3N h^2} \right)^{3/2} \right) + \frac{5}{2} \right]. \quad (2.10)$$

This very useful equation is called the **Sackur-Tetrode equation**. Increasing an ideal gas's volume, energy, or number of particles will increase the entropy.

Example 2.3. Volume dependence of entropy

For an ideal gas, if the volume of the gas increase from $V_i \rightarrow V_f$ in an isothermal process (U and N constant), the entropy changes by

$$\Delta S = Nk \ln \left(\frac{V_f}{V_i} \right).$$

From the isothermal expansion (1.3), the heat transferred into the system is

$$Q = NkT \ln \left(\frac{V_f}{V_i} \right).$$

This means

$$\Delta S = \frac{Q}{T}.$$

We will see that putting heat into a system will always increase its entropy in the next chapter.

Now consider another situation where we have a sudden, isolated expansion, which is also called **free expansion**, shown in Figure 2.7.

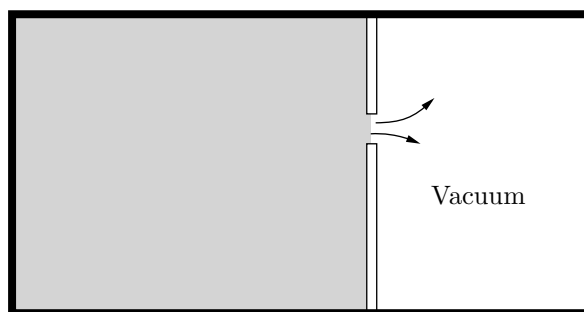


Figure 2.7: Free expansion of a gas into a vacuum.

The change in entropy in this process is still

$$\Delta S = Nk \ln \left(\frac{V_f}{V_i} \right).$$

Meanwhile, no work is done and no heat is transferred into the system, i.e. the internal energy remains constant. This means we can create entropy out of nothing.

2.6.2 Entropy of Mixing

Entropy is increased during a mixing process. Suppose we have two ideal gases of different species with the same volume, number of particles, and energy, separated by a partition. Now remove the partition. Entropy of both gases increases because their individual volume is doubled as they fill up the space. We can calculate the entropy change by treating each gas as a separated system, even after the mixing:

$$\Delta S_A = Nk \ln \frac{V_f}{V_i} = Nk \ln 2 = \Delta S_B.$$

Hence the total change is

$$\Delta S_{\text{total}} = \Delta S_A + \Delta S_B = 2Nk \ln 2.$$

This entropy increase has a name called the **entropy of mixing**.

What about the mixing of two ideal gases of the same species? We can view it in the way of adding gases. First let us look at the entropy change when a mole of argon (with same volume and energy) to a mole of helium. The entropy is approximately doubled:

$$S_{\text{total}} = S_{\text{He}} + S_{\text{Ar}} \approx 2S_{\text{He}}.$$

“Approximately” because the entropy of the argon is slightly greater than that of the helium due to its larger molecular mass. Now if a second mole of helium is added instead of argon, the entropy is not doubled. We can analyze the Sackur-Tetrode equation (2.10) and which parameters are changed during this process: both N and U are doubled, so U/N is unchanged,

$$S_i = Nk \left[\ln \left(\frac{V}{N} \left(\frac{4\pi m U}{3N h^2} \right)^{3/2} \right) + \frac{5}{2} \right] \longrightarrow S_f = (2N)k \left[\ln \left(\frac{V}{2N} \left(\frac{4\pi m U}{3N h^2} \right)^{3/2} \right) + \frac{5}{2} \right].$$

That is,

$$2S_i - S_f = 2Nk \ln 2$$

which is just the entropy of mixing. The difference between adding argon from adding helium is the result of the $1/N$ factor inside the logarithm in the Sackur-Tetrode equation. This factor comes from the derivation of this equation when we encounter $1/N!$ in the multiplicity function (2.8). Then what does this $1/N!$ factor mean in multiplicity? You would recall that this factor eliminates the double counting of microstates, due to the fact that molecules of the same species are *indistinguishable*. Argon atom and helium atom are distinguishable, but helium atom with another helium atom are not. If there were no such factor, the Sackur-Tetrode equation would become

$$S = Nk \left[\ln \left(V \left(\frac{4\pi m U}{3N h^2} \right)^{3/2} \right) + \frac{3}{2} \right].$$

J. Willard Gibbs noticed that something weird would happen (known as the **Gibbs paradox**) if this were the case.

3 INTERACTIONS AND IMPLICATIONS

3.1 Temperature

The second law discussed in the last chapter says that when two objects are in thermal equilibrium, their total entropy is at maximum. In chapter 1, we also say that two objects in thermal equilibrium have the same temperature. There surely is some connection between the entropy and temperature, as we will discuss in this section. Subsequently, we will find out what temperature truly is.

Consider a simple case, with two Einstein solids A and B sharing some total energy q_{total} . Figure 3.1 is a graph of the entropies S_A , S_B and S_{total} (some \ln functions of q for Einstein solids). At the equilibrium point, S_{total} is at maximum so we can write

$$\frac{\partial S_{\text{total}}}{\partial q_A} = 0 \quad \text{or} \quad \frac{\partial S_{\text{total}}}{\partial U_A} = 0 \quad [\text{at equilibrium}].$$

(Note that U_A is just some constants times q_A , which does not affect the result here.) The sum of the slopes of S_A and S_B is the slope of S_{total} . Hence,

$$\frac{\partial S_A}{\partial U_A} + \frac{\partial S_B}{\partial U_A} = 0 \quad [\text{at equilibrium}].$$

Since $U_A = U_{\text{total}} - U_B$, or $dU_A = -dU_B$, we can rewrite the equation as

$$\frac{\partial S_A}{\partial U_A} = \frac{\partial S_B}{\partial U_B} \quad [\text{at equilibrium}]. \quad (3.1)$$

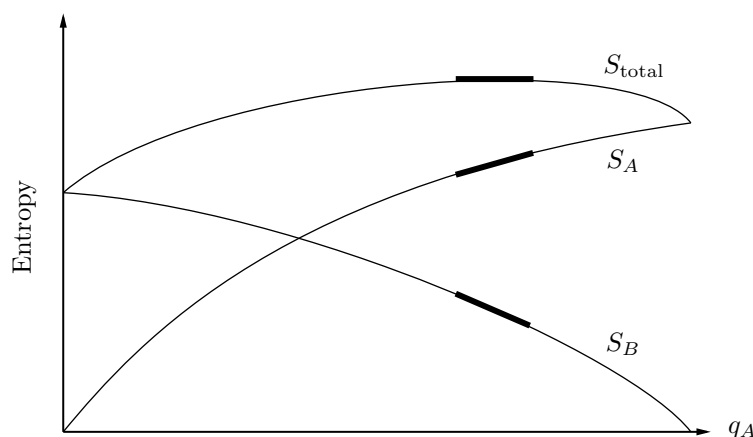


Figure 3.1: A plot of entropies vs q_A . The slopes of S_A and S_B are equal in magnitude when S_{total} is at maximum.

Heat energy flows into the object with the steeper $S(U)$ and out of the other one. The one with steeper $S(U)$ must have the lower temperature. Taking account for the correct units, these characteristics motivate us to write the definition of **temperature** to be

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_{N,V} \quad (3.2)$$

where we keep N and V constant.

Example 3.1. An Einstein solid

Recall that the entropy of an Einstein solid is

$$S = Nk \left[\ln \left(\frac{q}{N} \right) + 1 \right] = Nk \ln U - Nk \ln(\hbar\omega N) + Nk,$$

where we now consider the real energy $U = \hbar\omega q$ rather than energy units q . The temperature is

$$T = \left(\frac{\partial S}{\partial U} \right)^{-1} = \left(\frac{Nk}{U} \right)^{-1} \implies U = NkT.$$

This is exactly what we will get from the equipartition theorem (harmonic oscillators contribute $\frac{1}{2}kx^2$ and $\frac{1}{2}mv^2$). Note that N is the number of oscillators in the Einstein solid, not number of atoms. The number of atoms (in 3-dimensional space) should be $N/3$.

Example 3.2. A monatomic ideal gas

According to the Sackur-Tetrode equation (2.10), the entropy of a monatomic ideal gas is

$$S = Nk \ln V + Nk \ln U^{3/2} + (\text{a function of } N)$$

where N is the number of molecules. The temperature is

$$T = \left(\frac{\frac{3}{2}Nk}{U} \right)^{-1} \implies U = \frac{3}{2}NkT.$$

This is again consistent with the equipartition theorem.

3.2 Entropy and Heat

3.2.1 Predicting Heat Capacities

We already know the heat capacities of an Einstein solid with $q \gg N$,

$$C_V = \frac{\partial}{\partial T}(NkT) = Nk,$$

and for a monatomic ideal gas,

$$C_V = \frac{\partial}{\partial T} \left(\frac{3}{2}NkT \right) = \frac{3}{2}Nk.$$

These results agree with experimental results for low-density monatomic gases and for solids at reasonably high temperatures. Recalling what we did through whole chapter 2 and the last section, we can have a list of steps to predict the heat capacity of a system:

1. Use quantum mechanics and some combinatorics to find an expression for the multiplicity, Ω , in terms of U , V , N , and any other relevant variables.
2. Take the logarithm to find the entropy, S .
3. Do $\frac{\partial S}{\partial U}$ to find the temperature T as a function of U and other variables.
4. Solve for U as a function of T .
5. Do $\frac{\partial U}{\partial T}$ to predict the heat capacity with other variable fixed.

The most difficult step would be the first step. Later, Boltzmann statistics will help us find $U(T)$ without knowing the multiplicity or the entropy.

3.2.2 Measuring Entropies

Reversing steps 3-5 helps us measure the entropy of an system instead of calculate it theoretically. If we add amount of heat Q to a system with volume constant and no other forms of work, the entropy of the system changes by

$$dS = \frac{dU}{T} = \frac{Q}{T}.$$

This can be applied even when Q and dS are not infinitesimal when T is constant. However, when T is changing (but volume is not), we had better change the equation to

$$dS = \frac{C_V dT}{T}.$$

The change in entropy is then

$$\Delta S = \int_{T_i}^{T_f} \frac{C_V}{T} dT.$$

C_V usually is independent of temperature, but in some cases, C_V is a function of T at low T . We can calculate the total entropy if we know C_V all the way down to absolute zero,

$$S - S(0) = \int_0^{T_f} \frac{C_V}{T} dT.$$

The value of $S(0)$ is (in principle) zero because a system should be in its unique lowest-energy state at absolute zero: $\Omega = 1$ and $S = 0$. This is the **third law of thermodynamics**. Another way to present the third law of thermodynamics is that

$$C_V \rightarrow 0 \quad \text{as} \quad T \rightarrow 0$$

so that the integral from absolute zero to T_f does not diverge.

3.3 Paramagnetism

The **two-state paramagnet** is a system where equipartition theorem does not apply at all. We are to calculate its multiplicity and entropy.

3.3.1 Notation and Microscopic Physics

Consider a system with N spin-1/2 particles in a constant magnetic field \mathbf{B} in the $+z$ -direction. Each dipole (or particle) wants to align its magnetic dipole moment with the field. An **ideal paramagnet** is such a system if there are no interactions between dipoles. According to quantum mechanics, a particle's dipole moment is quantized. That is, a spin-1/2 particle can only be spin up “ \uparrow ”, or spin down “ \downarrow ” along the z -axis. Let us define the energy of a \uparrow particle to be $-\mu B$ and the energy of a \downarrow particle to be μB , where μ is a constant related to the particle's magnetic moment. The amount of energy required to “flip” a dipole from up to down is thus $2\mu B$. The total energy of the system is

$$U = \mu B(N_{\downarrow} - N_{\uparrow}) = \mu B(N - 2N_{\uparrow}), \quad (3.3)$$

where N_{\uparrow} and N_{\downarrow} are the numbers of up and down dipoles respectively, and $N = N_{\uparrow} + N_{\downarrow}$. Define the **magnetization**, M , to be the total magnetic moment of the whole system,

$$M = \mu(N_{\uparrow} - N_{\downarrow}) = -\frac{U}{B}. \quad (3.4)$$

The goal is to determine how U and M depends on temperature. Just like flipping coins, a two-state paramagnet is an Ising system. The multiplicity of a paramagnet with N_{\uparrow} up dipole is simply

$$\Omega(N_{\uparrow}) = \binom{N}{N_{\uparrow}} = \frac{N!}{N_{\uparrow}!N_{\downarrow}!}. \quad (3.5)$$

3.3.2 Numerical Solution

Suppose there is a paramagnet which contains 100 dipoles. We can numerically calculate its multiplicity and then the entropy by taking the logarithm. Figure 3.2 is a plot of entropy vs energy.

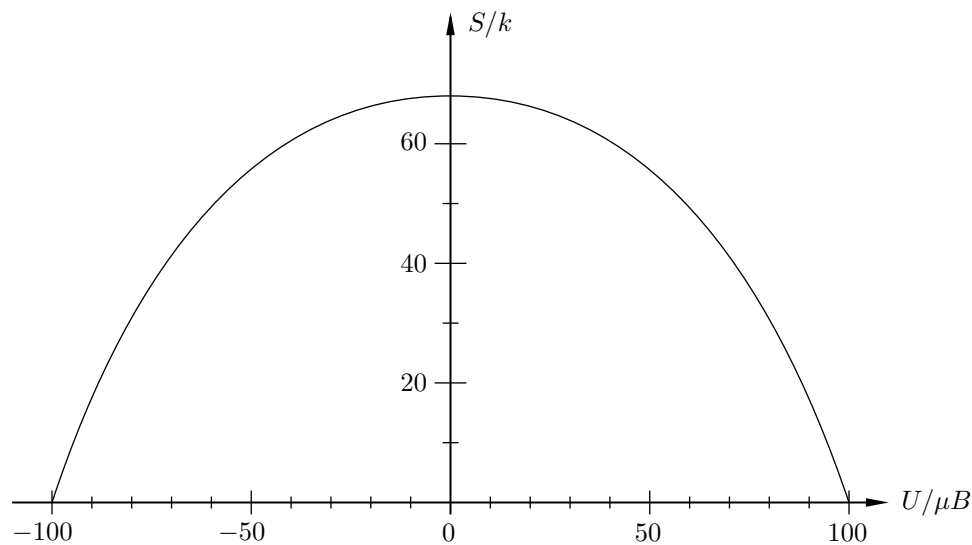


Figure 3.2: Entropy as a function of energy for a two-state paramagnet consisting of 100 elementary dipoles.

The largest entropy is at $U = 0$, the point where exactly half of the dipoles point down. Before this point, the system tends to absorb energy to increase its entropy, the same behavior as an Einstein solid or an ideal gas. However, after reaching $U = 0$, adding energy to the system decreases its entropy. That is, it will spontaneously give up energy to any nearby object whose entropy-energy graph has a positive slope.

Now let us discuss about temperature. When $U < 0$, the temperature of the paramagnet is positive. Weird things happen when $U \geq 0$. First, the temperature is infinite at $U = 0$ because the slope vanishes. This means the system will give energy to any other system whose temperature is finite. Second, the temperature at $U > 0$ is negative according to the definition of temperature. Such “negative” temperature behaves as if it is higher than positive temperatures: a system with negative temperature will transfer energy to any system with positive temperature. The study of unusual behaviors of paramagnet temperature gives us a hint that entropy is the more fundamental quantity of thermodynamics. Temperature is just the “willingness” to give up energy, a intermediate character between energy and entropy.

3.3.3 Analytic Solution

If we want to study the relationship between multiplicity and temperature or magnetization, it is convenient to calculate some general formulas. For simplicity, let us directly calculate the entropy from (3.5) using Stirling’s approximation.

$$\begin{aligned} S/k &= \ln N! - \ln N_{\uparrow}! - \ln(N - N_{\uparrow})! \\ &\approx N \ln N - N - N_{\uparrow} \ln N_{\uparrow} + N_{\uparrow} - (N - N_{\uparrow}) \ln(N - N_{\uparrow}) + (N - N_{\uparrow}) \\ &= N \ln N - N_{\uparrow} \ln N_{\uparrow} - (N - N_{\uparrow}) \ln(N - N_{\uparrow}) \end{aligned}$$

where we assume that N , N_{\uparrow} and $N_{\downarrow} = N - N_{\uparrow}$ are large numbers. The temperature of the paramagnet is given by

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_{N,B} = \frac{\partial N_{\uparrow}}{\partial U} \frac{\partial S}{\partial N_{\uparrow}} = -\frac{1}{2\mu B} \frac{\partial S}{\partial N_{\uparrow}}.$$

After some algebra calculating $\partial S / \partial N_{\uparrow}$, we arrive at

$$\frac{1}{T} = \frac{k}{2\mu B} \ln \left(\frac{N - U/\mu B}{N + U/\mu B} \right). \quad (3.6)$$

Solving for U in terms of T ,

$$U = N\mu B \left(\frac{1 - e^{2\mu B/kT}}{1 + e^{2\mu B/kT}} \right) = -N\mu B \tanh \left(\frac{\mu B}{kT} \right), \quad (3.7)$$

The resultant magnetization is

$$M = N\mu \tanh \left(\frac{\mu B}{kT} \right) \quad (3.8)$$

If you look at the graph of \tanh , you will find that at very small temperature the system is completely magnetized. The magnetization goes to zero if $T \rightarrow \infty$. Negative magnetization gives a negative temperature.

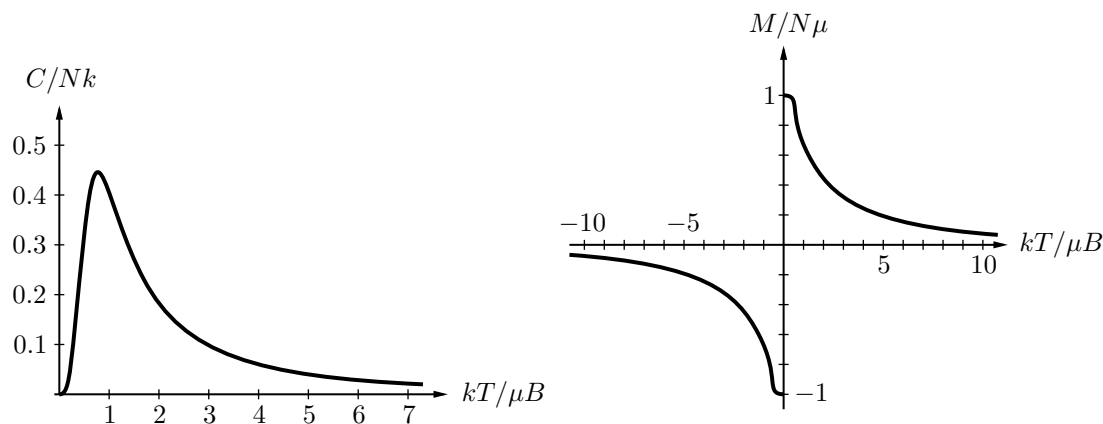


Figure 3.3: Heat capacity and magnetization of a two-state paramagnet.

The heat capacity of the paramagnet is (from 3.7)

$$C_B = \left(\frac{\partial U}{\partial T} \right)_{N,B} = Nk \cdot \frac{(\mu B/kT)^2}{\cosh^2(\mu B/kT)}. \quad (3.9)$$

At both $T = 0$ and $T \rightarrow \infty$, the heat capacity goes to zero. Figure 3.3 are some graphs related to discussions above. Once we get U as a function of temperature, we can do the same for entropy (see Appendix A.1.4),

$$S = Nk[\ln(2 \cosh x) - x \tanh x], \quad \text{where } x = \mu B/kT. \quad (3.10)$$

3.4 Mechanical Equilibrium and Pressure

In previous sections we focused on the exchange of energy between two systems. Systems in thermal equilibrium have the same temperature. There are other exchange quantities and corresponding equilibriums. In this section, we will talk about the exchange of volume, governed by the pressure of two systems. When two systems have the same pressure, we say that they are in **mechanical equilibrium**.

Consider two systems, which are isolated from other systems, separated by a movable partition that allows them to exchange both energy and volume (see Figure 3.4).

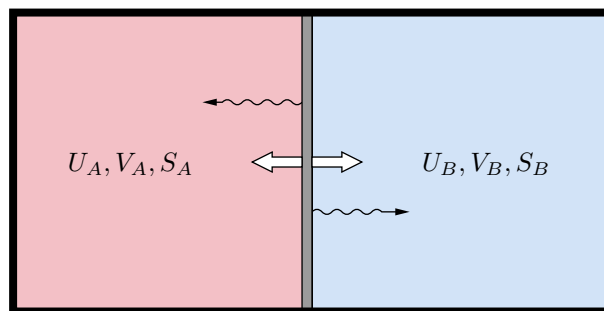


Figure 3.4: Two systems that can exchange both energy and volume with each other. The total energy and volume are fixed.

In this scenario, we have conserved energy and volume,

$$U_A + U_B = U_{\text{total}} \quad \text{and} \quad V_A + V_B = V_{\text{total}}.$$

The entropy now is a function of both U_A and V_A . The equilibrium occurs where S_{total} is at maximum, at which

$$\frac{\partial S_{\text{total}}}{\partial U_A} = 0, \quad \frac{\partial S_{\text{total}}}{\partial V_A} = 0.$$

The first equality says that the two systems are at the same temperature, so let us study the second equality. Similar to what we have done in the definition of temperature, we have

$$0 = \frac{\partial S_{\text{total}}}{\partial V_A} = \frac{\partial S_A}{\partial V_A} + \frac{\partial S_B}{\partial V_A} = \frac{\partial S_A}{\partial V_A} - \frac{\partial S_B}{\partial V_B}.$$

Therefore,

$$\frac{\partial S_A}{\partial V_A} = \frac{\partial S_B}{\partial V_B} \quad [\text{at equilibrium}].$$

We know that the pressure of the two system are now the same because they are at mechanical equilibrium, so it is reasonable to guess that pressure can be written in forms of $\partial S/\partial V$. There are two ways to think about it: 1. When $\partial S/\partial V$ is large, it means the system gains a lot of entropy by expanding a little. This tendency of increasing entropy says that the system wants to expand, hence the pressure is large. Thus, pressure should be proportional to $\partial S/\partial V$. 2. The dimensional analysis says that the pressure should have the same units as $T(\partial S/\partial V)$. Thus, we shall propose the relation between pressure and entropy to be

$$P = T \left(\frac{\partial S}{\partial V} \right)_{U,N} \quad (3.11)$$

Example 3.3. Pressure of a monatomic ideal gas

We can use the ideal gas to check whether (3.11) is correct. The Sackur-Tetrode equation (2.10) says that the entropy of a monatomic gas is

$$S = Nk \ln V + \frac{3}{2} Nk \ln U + k \ln f(N).$$

The pressure of the gas should be

$$P = T \left(\frac{\partial S}{\partial V} \right) = T \frac{\partial}{\partial V} (Nk \ln V) = \frac{NkT}{V},$$

or

$$PV = NkT.$$

This is the ideal gas law. It shows that no constant factors should be added to (3.11).

3.4.1 The Thermodynamic Identity

There is an equation that summarizes what we have learnt in this chapter about entropy, temperature, and pressure. By the chain rule, an infinitesimal change in entropy can be written as

$$dS = \left(\frac{\partial S}{\partial U} \right)_V dU + \left(\frac{\partial S}{\partial V} \right)_U dV = \frac{1}{T} dU + \frac{P}{T} dV.$$

After some rearrangement,

$$dU = T dS - P dV. \quad (3.12)$$

This equation is called the **thermodynamic identity**, provided that T and P are well defined and no other relevant variables (such as number of particles) are changing.

3.5 Diffusive Equilibrium and Chemical Potential

What happens if we now allow particles and energy to exchange through a membrane between two systems? Again there are some

similar arguments: at the maximum of total entropy,

$$\left(\frac{\partial S_{\text{total}}}{\partial U_A}\right)_{N_A, V_A} = 0, \quad \text{and} \quad \left(\frac{\partial S_{\text{total}}}{\partial N_A}\right)_{U_A, V_A} = 0.$$

The second equation says that

$$\frac{\partial S_A}{\partial N_A} = \frac{\partial S_B}{\partial N_B} \quad [\text{at equilibrium}].$$

By convention, we multiply a $-T$ on both sides to obtain

$$-T \frac{\partial S_A}{\partial N_A} = -T \frac{\partial S_B}{\partial N_B} \quad [\text{at equilibrium}].$$

The quantities on both sides are called the **chemical potential**, denoted by μ ,

$$\mu \equiv -T \left(\frac{\partial S}{\partial N}\right)_{U, V} \quad (3.13)$$

When two systems are in **diffusive equilibrium**, they are at the same chemical potential. Particles tend to flow from the system with higher μ to the system with lower μ (which is a usual way to define a “potential”; this fact comes from the minus sign in $-T$). For example, a monatomic ideal gas will have a chemical potential (by differentiating the Sackur-Tetrode equation with respect to N)

$$\mu = -kT \ln \left[\frac{V}{N} \left(\frac{2\pi m kT}{h^2} \right)^{3/2} \right] \quad (3.14)$$

where we replace $U = \frac{3}{2}NkT$. Higher chemical potential corresponds to more molecules when other parameters are held constant. Thus, molecules in ideal gases want to move to places of less molecules—a sign of diffusion, which is part of the name “diffusive equilibrium”.

With the idea of chemical potential, we can generalize the thermodynamic identity to

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial U}\right)_{N, V} dU + \left(\frac{\partial S}{\partial V}\right)_{N, U} dV + \left(\frac{\partial S}{\partial N}\right)_{U, V} dN \\ &= \frac{1}{T} dU + \frac{P}{T} dV - \frac{\mu}{T} dN. \end{aligned}$$

Rearranging some terms,

$$dU = T dS - P dV + \mu dN. \quad (3.15)$$

Just as the $-P dV$ is usually related to mechanical work, the μdN is often referred to as “chemical work”. Note that μ has units of energy.

4 ENGINES AND REFRIGERATORS

I do not like this chapter, so the notes are somewhat plain and simple.

4.1 Heat Engines

A **heat engine** is a device which undergoes cyclic processes and converts heat into work. Only *part* of the heat can be converted to work. This is because entropy is also flowing into the engine, and to dispose the entropy the heat engine must transfer some heat to the environment. One attribute of the heat engine is that it returns to the initial condition after each cycle, same temperature, same entropy, etc.

To make a heat engine work, amount of heat Q_h comes from a **hot reservoir** and flows into a **working substance** (like an ideal gas) that does work W , which later cools down by heat flowing to a **cold reservoir** of the amount Q_c . A reservoir is assumed to be *big* so that its temperature is constant. Figure 4.1 is an energy flow diagram of heat engines. Note that in this chapter, Q_c , Q_h , and W are all positive. The first law of thermodynamics requires that

$$Q_h = W + Q_c.$$

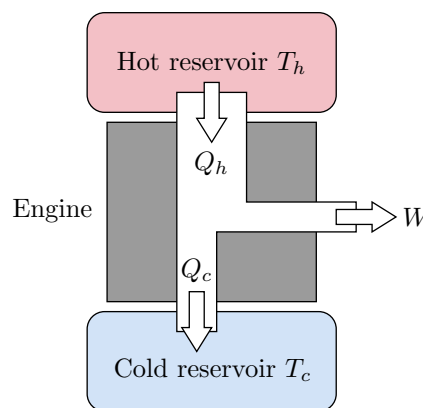


Figure 4.1: Energy-flow diagram for a heat engine.

4.1.1 Efficiency

The efficiency of a heat engine is defined by

$$e \equiv \frac{W}{Q_h} = 1 - \frac{Q_c}{Q_h}. \quad (4.1)$$

We have mentioned that the entropy of the engine must not accumulate after each cycle, so the entropy it expels must be at least the entropy it absorbs, $S_{\text{out}} \geq S_{\text{in}}$, or equivalently

$$\frac{Q_c}{T_c} \geq \frac{Q_h}{T_h} \implies \frac{Q_c}{Q_h} \geq \frac{T_c}{T_h}$$

provided that the temperature of reservoirs does not change. This means that the efficiency has some limit:

$$e \leq 1 - \frac{T_c}{T_h} \quad (4.2)$$

4.1.2 Carnot Cycle

A **Carnot cycle** achieves the limit of the efficiency, i.e., no extra entropy is produced during the cycle. To minimize the entropy creation, the temperature of the engine T_e must be close to T_h ,

$$T_e \lesssim T_h,$$

because in this way, $S_h = Q/T_h \sim S_e = Q/T_e$. We also want T_e to remain constant during the heat flowing. This means the engine is undergoing an isothermal process with heat flowing in, or isothermal expansion. Similarly when heat is flowing from the engine to the cold reservoir, their temperature must be close, $T_e \gtrsim T_c$. This time the heat engine is undergoing an isothermal compression. In between, T_e changes from T_h with T_c with no heat flows so that there is no entropy created. Thus they are adiabatic processes, specifically, adiabatic expansion from T_h to T_c and adiabatic compression from T_c to T_h .

The work done by the engine is the area under both P - V diagram and T - S diagram (shown in Figure 4.2). By thermodynamic identity (3.12),

$$\oint dU = \oint T dS - \oint P dV = 0 \implies W = \oint P dV = \oint T dS.$$

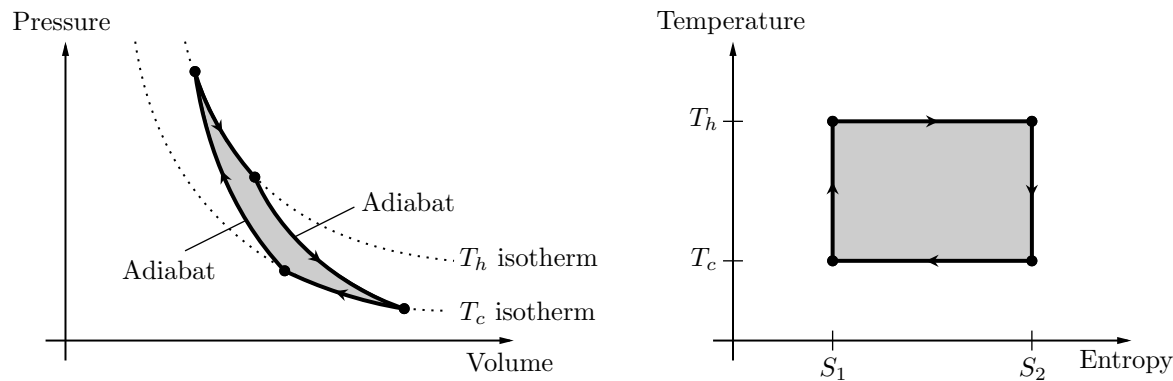


Figure 4.2: P - V diagram and T - S diagram for a monatomic ideal gas undergoing a Carnot cycle.

Note that the entropy produced by the Carnot cycle is zero, so it is reversible. For any irreversible heat engines, the net entropy produced after a cycle is not zero, but rather

$$\Delta S_{\text{res}} = \oint \frac{dQ}{T} \geq 0.$$

This is called the **Clausius inequality**. A Carnot cycle is maximally efficient, but it is *impractical*. This is because the heat transfer between the hot reservoir and the engine takes forever if the two temperatures are close to each other. It is the same when the engine is exchanging heat with the cold reservoir. Therefore, a Carnot engine is only a *theoretical* engine. In real life, we need to balance the efficiency with time, or more precisely, the power of the engine.

4.1.3 Another Example: Otto Cycle

The **internal combustion engines** are found in most automobiles and it relies on a cycle called the **Otto cycle**. The working substance is a gas with heat capacity C_V . Figure 4.3 shows a cyclic diagram for the Otto cycle.

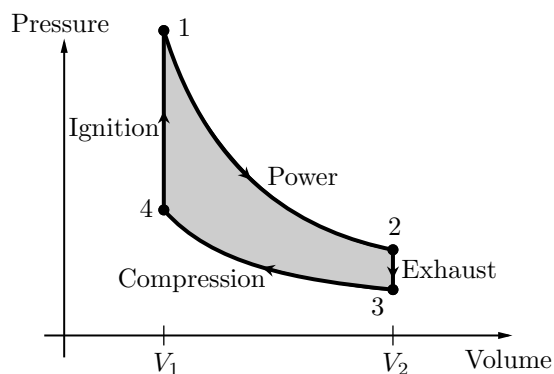


Figure 4.3: The idealized Otto cycle, an approximation of what happens in a gasoline engine.

The Otto engine is an irreversible engine. The mixture of air and vaporized gasoline is first injected into a cylinder and compressed adiabatically by a piston (3-4). A spark plug then ignites the mixture, raising its temperature and pressure while the volume does

not change (4-1). The high-pressure gas pushes the piston outward, expanding adiabatically and producing mechanical work (1-2). Finally, the hot exhaust gases are expelled and replaced by a new mixture at lower temperature and pressure (2-3).

The heat flowing in from the hot reservoir and the heat flowing out to the cold reservoir are respectively

$$Q_h = \int_{T_4}^{T_h} C_V dT = C_V(T_h - T_4), \quad \text{and similarly} \quad Q_c = C_V(T_2 - T_c),$$

where $T_h = T_1$ and $T_c = T_3$. In the two adiabatic processes, using $\gamma = (f + 2)/f$,

$$T_2 V_2^{\gamma-1} = T_h V_1^{\gamma-1} \implies T_2 = T_h \left(\frac{V_1}{V_2} \right)^{\gamma-1},$$

and likewise,

$$T_4 = T_c \left(\frac{V_c}{V_4} \right)^{\gamma-1}.$$

This gives the relation

$$T_2 T_4 = T_h T_c.$$

The efficiency turns out to be

$$e = 1 - \frac{Q_c}{Q_h} = 1 - \frac{C_V(T_2 - T_c)}{C_V(T_h - T_4)} = 1 - \left(\frac{V_1}{V_2} \right)^{\gamma-1} = 1 - \frac{T_c}{T_h} \cdot \frac{T_2}{T_4}.$$

We see that the efficiency depends on the ratio of the volumes, which is called the **compression ratio**.

4.2 Refrigerators

4.2.1 Definition

A **refrigerator** is a heat engine which extract heat from the cold reservoir to a hot reservoir (see Figure 4.4).

The “efficiency”, often called the **coefficient of performance**, is defined as

$$\text{COP} \equiv \frac{Q_c}{W} = \frac{1}{Q_h/Q_c - 1}. \quad (4.3)$$

By the second law, $S_h \geq S_c$, which means

$$\frac{Q_h}{T_h} \geq \frac{Q_c}{T_c} \implies \frac{Q_h}{Q_c} \geq \frac{T_h}{T_c}.$$

The efficiency has a limit

$$\text{COP} \leq \frac{1}{T_h/T_c - 1} = \frac{T_c}{T_h - T_c} \quad (4.4)$$

Note that the efficiency of refrigerators can be greater than 1.

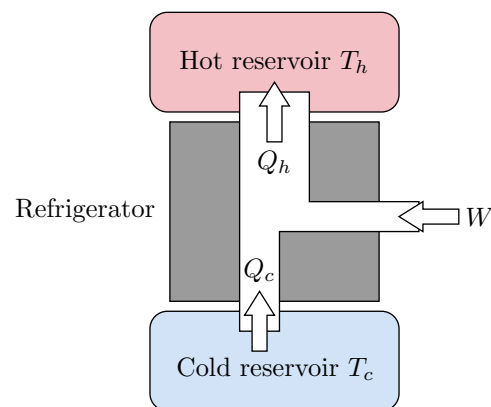


Figure 4.4: Energy-flow diagram for a refrigerator or air conditioner.

4.2.2 Second Law Revisited

Heat engines and refrigerators promotes not only industrial developments, but also the understanding of the second law of thermodynamics and entropy. Both Clausius and Kelvin stated the second law in terms of heat engines. **Clausius statement** of the second law is: no process is possible whose sole result is the transfer of heat from a colder to a hotter body. Simply, refrigerators need to be plugged in. **Kelvin statement** of the second law is: no process is possible whose sole result is the complete conversion of heat into work. Simply, engines produce waste heat.

4.3 Decade Thermometer

The decade thermometer (Table 4.1) shows the temperature in Kelvin scale by order of magnitudes. We will discuss how to cool things down and what will happen at very low temperatures.

Scale	Temperature	Item
10^{12}	5.5 TK	Highest man-made temperature in thermal equilibrium at LHC
10^{10}	10 GK	Supernovae
10^7	15.6 MK	Sun's core temperature
10^3	5780 K	Sun's surface temperature
	3915 K	Sublimation point of carbon
	3695 K	Melting point of tungsten
	1811 K	Melting point of iron
10^2	293 K	Room temperature
10^1	77.35 K	Boiling point of nitrogen
1	4.22 K	Boiling point of helium
	4.1 K	Superconductivity point of mercury
	2.725 K	Cosmic microwave background
10^{-2}	60 mK	Adiabatic demagnetization of paramagnetic molecules
10^{-3}	2.5 mK	Helium-3 exhibits superfluidity
	1.7 mK	temperature record for helium-3/helium-4 dilution refrigeration
10^{-6}	1 μ K	Nuclear demagnetization and laser cooling
10^{-9}	1 nK	Bose-Einstein condensation
10^{-11}	38 pK	lowest laboratory-produced temperature

Table 4.1: Decade Thermometer

4.3.1 Isentropic Expansion

The simplest way to cool down gases is to let it do work against some force. For an ideal gas, PV^γ is constant, or equivalently, $TV^{\gamma-1}$, in an isentropic (or simply adiabatic) expansion, where $\gamma = 5/3$. From the first relation, we get

$$P \left(\frac{T}{P} \right)^\gamma = \text{const.} \implies T \propto P^{2/5}.$$

From the second relation, we get

$$T \propto V^{1-\gamma} = V^{-2/3}.$$

If we can decrease the pressure or increase the volume of an ideal gas, we can cool down a gas.

4.3.2 The Throttling Process

The **throttling** process is also known as the **Joule-Thomson process**. Basically, suppose a piston is exerting a pressure P_i on a fluid on the left side. Another piston, exerting pressure P_f , on the other (right) side of the fluid is moving backward on the other side to make room, and $P_f < P_i$. Let the initial and final volume be V_i and V_f , respectively. There is no heat flow in this process, so the change in the energy of the fluid is

$$U_f - U_i = Q + W = W_{\text{left}} + W_{\text{right}} = P_i V_i - P_f V_f.$$

Rearranging the terms,

$$U_f + P_f V_f = U_i + P_i V_i \quad \text{or} \quad H_f = H_i.$$

The enthalpy is constant during the throttling process. However, for an ideal gas, constant enthalpy means constant temperature:

$$H = U + PV = \frac{f}{2} NkT + NkT = \frac{f+2}{2} NkT.$$

There will be no cooling at all. Fortunately, real gases are not ideal. Recall that ideal gases assume that intermolecular forces are negligible, but real gas has a force between molecules: weakly attractive at long distances, and strongly repulsive at short distances. Thus, the real enthalpy is

$$H = U_{\text{potential}} + U_{\text{kinetic}} + PV.$$

At large distance, the attraction force give rise to a negative potential, and it is less negative as the pressure decreases and the intermolecular distances increases. An increase in potential energy means a decrease in kinetic energy, and hence the temperature.

4.3.3 Liquefaction of Gases

In 1895 William Hampson and Carl von Linde independently invented a device that can liquefy oxygen or nitrogen. The compressed gas is first cooled (to room temperature is sufficient for nitrogen or oxygen) and then passed through a heat exchanger on its way to a throttling valve. The gas cools upon throttling and returns through the heat exchanger to further cool the incoming gas. Eventually the incoming gas becomes cold enough to partially liquefy upon throttling. New gas must be added at the compressor to replace what is liquefied.

4.3.4 Pumping

Suppose we now have some liquefied helium in a container. We can pump out the gas in the container to reduce the pressure until it is below the equilibrium vapor pressure so that liquid helium evaporates. Latent heat of the liquid-gas phase transition is removed along with the vapor. Repeating this process can reduce the temperature to fractions of a Kelvin.

4.3.5 Dilution Refrigerator

The **helium dilution refrigerator** can get helium from 1 K to several mK. At subkelvin temperature scales, two isotopes of helium (^4He and ^3He) are relatively immiscible. Typically, ^3He is above ^4He because it is lighter. Below about 0.1 K, no ^4He will dissolve in ^3He , but about 6% of ^3He will dissolve into ^4He . If we remove the ^3He atoms dissolved in ^4He and let it evaporate, other ^3He atoms will dissolve in ^4He liquid so that the 6% is met. This process is a heat absorbing process. In this case, ^3He is the working substance (it will circulate back by some process), and ^4He is cooled.

4.3.6 Adiabatic Demagnetization

The **Adiabatic demagnetization** is another way to get temperatures below 10 mK. Recall that the magnetization of a paramagnet is

$$M = \mu(N_{\uparrow} - N_{\downarrow}) = N\mu \tanh\left(\frac{\mu B}{kT}\right).$$

It is directly related to the number of dipoles pointing up, and so directly related to entropy. Suppose we have an electronic paramagnet, where its majority of the dipoles are pointing up initially. Lowering the external magnetic field can reduce its temperature linearly if the magnetization remains constant. This is achieved when the process is adiabatic where entropy does not change.

The demagnetization refrigerator involves a magnet and a large reservoir. The reservoir is in thermal contact with the magnet so that it remains at a constant temperature. If we increase the magnetic field, the entropy in the magnet decreases. Now remove the reservoir and lower the magnetic field (an adiabatic process). The entropy does not change, but the temperature move from the large magnetic field curve to small field curve (see Figure 4.5). The figure shows that at high temperature, the dipole becomes random and the entropy reaches a constant value. With higher field strength, the entropy rises gradually rather than suddenly with temperature because the dipoles tend to remain aligned with the field.

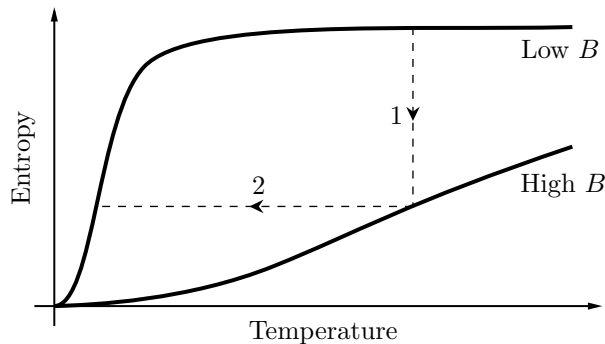


Figure 4.5: Entropy as a function of temperature for an ideal two-state paramagnet, at two different values of magnetic field strength. The magnetic cooling process consists of an isothermal increase in the field strength (step 1), followed by an adiabatic decrease (step 2).

However, it seems like changing the magnetic field has not done anything useful, we are just saying the temperature changes in the magnet but it does not have any quadratic degree's of freedom. Thus, a solid (maybe an Einstein solid) with higher temperature is attached to the magnet so that they are in thermal contact. In this way we can *physically* cool down a solid. At sufficiently low temperature, the paramagnet becomes not ideal. Even if we set the external magnetic field to zero, there will still be fields from the dipoles themselves and they mutually interact with each other. Hence it is not possible to reach absolute zero by just eliminating the external magnetic field. For electronic paramagnets (with relatively strong dipole-dipole interactions), the cooling limit is about 1 mK. For a *nuclear* paramagnet with much weaker interactions, the lowest temperature produced is 2.8×10^{-10} K.

4.3.7 Laser Cooling

Consider one rubidium atom (which can be seen as a “low density” gas) with a laser shining on it. The laser emits photons coherently with momentum $\hbar k$ where k is the wavenumber of the photon. The momentum of the rubidium atom changes after it absorbs the photon,

$$p_{\text{Rb}} \rightarrow p_{\text{Rb}} + \hbar k,$$

if the incident photon is at the right excitation wavelength (about 780 nm). After some small amount of time, the atom undergoes a spontaneous emission, but in random direction. On average, large number of spontaneous emissions contribute to a zero net force on the atom because of the randomness. The change in velocity of the atom is then only due to the force from the laser beam,

$$\Delta v = \frac{\hbar k}{m} = v_{\text{recoil}}.$$

If there are Γ spontaneous emission per second, the total force on the atom is approximately

$$F = m \frac{dv}{dt} = m \frac{\hbar k}{m} \Gamma = \hbar k \Gamma.$$

How are we supposed to cool down a gas with a force? A low density gas will have a velocity distribution, and we already know that smaller average speed means a lower temperature. If we apply a uniform force on the gas, they will move uniformly to some direction and nothing is done to temperature. This indicates that the force must be velocity dependent. The simplest tool to make such a force is by using Doppler shifts.

Suppose now the laser beam is set at a lower frequency. An atom at rest with respect to the laser rarely absorbs the photon, but an atom moving towards the laser does if it has the right velocity so that the photon is blue-shifted to the right frequency. The average force is thus highest at such velocity while lower at others (see left diagram of Figure 4.6). Atoms moving opposite to the laser beam will see red-shifted photons, so they do not absorb photons either. To solve this, we simply place another laser on the other side (in fact, 6 lasers are needed in three dimensions). The force-velocity relation is shown in the right diagram of Figure 4.6. For small v , the force is approximated as

$$F = -2\alpha v.$$

The gas atoms is immersed in a “space with viscosity”, reducing their speeds.

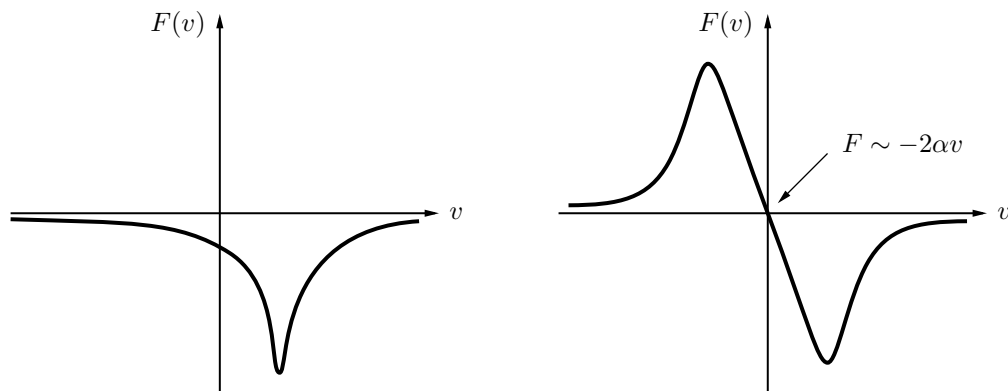


Figure 4.6: Force-velocity diagram of one laser (left) and two opposite lasers (right). When there are two opposite lasers, the force-velocity is linear for small v , similar to a drag force.

At very low temperature, the recoil of the atom from spontaneous emission can no longer be averaged out. This sets a limit to laser cooling, it starts to fail at the Doppler temperature

$$T_{\text{Doppler}} \sim \frac{\hbar\Gamma}{2}.$$

For rubidium, this temperature is about $140 \mu\text{K}$.

Some Comments

Why should we try to lower the temperature of substances? It turns out that at fractions of a Kelvin, helium starts to exhibit superfluidity. At even lower temperatures like mK, μK , or even nK, there are more interesting things happening, such as magnetic behaviors of atoms and nuclei, and Bose-Einstein condensations.

5 FREE ENERGY

There are many thermodynamical processes, such as chemical reactions, that are not cyclic like heat engines. This chapter applies laws of thermodynamics to chemical reactions and other transformations of matter. These processes may not be isolated—they may interact with environments thermally and mechanically. Fixed and unfixed parameters are also different. For example, energy and volume is often unfixed, but temperature and pressure are fixed. There are new quantities appearing that are useful to describe constant-temperature and constant-pressure processes.

5.1 Free Energy as Available Work

5.1.1 Thermodynamic Potentials

Entropy is a “function of state”, which means it is path independent. That is, this physical quantity does not depend on how you get to the state. The work W is in general not a function of state: as you may observed, the area under a P - V graph clearly depends on the path taken from one state to another (work done by conservative forces like gravity is in fact path independent). Other functions of state can be the internal energy U , all gas quantities, P , V , T , and combinations of them. For example, the enthalpy is also a function of state,

$$H = U + PV.$$

It is the total energy needed to create a system out of nothing. Alternatively, it can be interpreted as the energy you gain by completely annihilating the system. There are two useful combinations of those gas quantities,

$$F \equiv U - TS, \quad \text{and} \quad G \equiv U + PV - TS,$$

where F is the **Helmholtz free energy** and G is the **Gibbs free energy**. The Helmholtz free energy is useful in constant-temperature environment. It is the total energy needed to created the system, minus the heat you can get for free from an environment at temperature T . Thus, F is the work needed to created the system. The Gibbs free energy, is the Helmholtz free energy (the work to create the system), plus the work you need to do to make a space for it under atmospheric pressure. The four functions, U , H , F and G , are called **thermodynamic potentials**.

Example 5.1. Relations between thermodynamic potentials

Suppose you want to create a bear with volume V and total internal energy U , under an environment with atmospheric temperature T . The work you need to create the bear *itself* is not U , but the Helmholtz free energy $F = U - TS$. You do not need to summon the entire energy, there will be some heat TS from the environment that helps you create the bear.

Now you have already create the bear (maybe in vacuum with the help of some mysterious heat), and now you want to make a room for it in some atmosphere with pressure P . The work you need to create the room against the atmosphere is PV . Thus the total work necessary is the work creating the bear, plus the work creating the room, is the Gibbs free energy $G = F + PV = U - TS + PV$. Finally, the total energy for creating a bear and make it some room, from you, and from the environment, is the enthalpy: $H = U + PV = G + TS$.

In general, we would like to study the changes in these energies. At constant temperature environment,

$$\Delta F = \Delta U - T\Delta S = Q + W - T\Delta S.$$

If all entropy comes from the heat added Q , then $Q = T\Delta S$ and the change in F is exactly the work done on the system. No new entropy is created in this process, so it is reversible. If some entropy is created not by the heat, then $Q < T\Delta S$, so

$$\Delta F \leq W \quad \text{at constant } T.$$

At constant temperature and pressure environment, we should think about the Gibbs free energy:

$$\Delta G = \Delta U - T\Delta S + P\Delta V = Q + W - T\Delta S + P\Delta V.$$

Here work W includes the work done by the environment and other work,

$$W = -P\Delta V + W_{\text{other}}.$$

Again with $Q - T\Delta S \leq 0$, we have

$$\Delta G \leq W_{\text{other}} \quad \text{at constant } T, P.$$

5.1.2 Thermodynamic Identities

Recall that the thermodynamic identity (3.15) says

$$dU = T dS - P dV + \mu dN.$$

However, U may not be the most useful energy in some situation. For example, the free variable of U contains the entropy S , which is usually uncontrollable in experiments. We need to work with other thermodynamic potentials such as H , F , or G for more controllable variables.

The definition of the Helmholtz energy $F \equiv U - TS$ says that

$$dF = dU - T dS - S dT = T dS - P dV + \mu dN - T dS - S dT$$

where we use the chain rule and the thermodynamic identity of U . This gives the “thermodynamic identity for F ”:

$$dF = -S dT - P dV + \mu dN. \quad (5.1)$$

We see that by subtracting a TS term from U , the variable in the original thermodynamic identity changes from S to T . Such a change is called **Legendre transformation**. Now the entropy S is not a free variable, it can be determined (along with other quantities)

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

Similarly, we can find the “thermodynamic identity for G ”

$$dG = -S dT + V dP + \mu dN \quad (5.2)$$

with

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

The “thermodynamic identity for H ” is

$$dH = T dS + V dP + \mu dN \quad (5.3)$$

Table 5.1 summarize the discussion above.

Potentials	Differentials	Free variables	First derivatives	
U	$dU = T dS - P dV + \mu dN$	S, V, N	$T = \left(\frac{\partial U}{\partial S}\right)_{V,N}$	$P = -\left(\frac{\partial U}{\partial V}\right)_{S,N}$
$H = U + PV$	$dH = T dS + V dP + \mu dN$	S, P, N	$T = \left(\frac{\partial H}{\partial S}\right)_{P,N}$	$V = \left(\frac{\partial H}{\partial P}\right)_{S,N}$
$F = U - TS$	$dF = -S dT - P dV + \mu dN$	T, V, N	$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N}$	$P = -\left(\frac{\partial F}{\partial V}\right)_{T,N}$
$G = H - TS$	$dG = -S dT + V dP + \mu dN$	T, P, N	$S = -\left(\frac{\partial G}{\partial T}\right)_{P,N}$	$V = \left(\frac{\partial G}{\partial P}\right)_{T,N}$

Table 5.1: Four thermodynamic potentials, their differentials, free variables, and first derivatives (except chemical potentials).

5.1.3 Second Derivatives and Maxwell Relations

Suppose all thermodynamic potentials are smooth so that their second derivatives commute: $\partial_{xy} = \partial_{yx}$. This is also saying that

$$F_x = \left(\frac{\partial f}{\partial x} \right)_y, \quad F_y = \left(\frac{\partial f}{\partial y} \right)_x \implies \frac{\partial F_y}{\partial x} = \frac{\partial F_x}{\partial y}.$$

Then we can get **Maxwell relations** by taking derivatives to the first derivatives:

$$T = \left(\frac{\partial U}{\partial S} \right)_{V,N}, \quad P = - \left(\frac{\partial U}{\partial V} \right)_{S,N} \implies \left(\frac{\partial T}{\partial V} \right)_S = - \left(\frac{\partial P}{\partial S} \right)_V.$$

$$T = \left(\frac{\partial H}{\partial S} \right)_{P,N}, \quad V = \left(\frac{\partial H}{\partial P} \right)_{S,N} \implies \left(\frac{\partial T}{\partial P} \right)_S = \left(\frac{\partial V}{\partial S} \right)_P.$$

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

$$S = - \left(\frac{\partial G}{\partial T} \right)_{P,N}, \quad V = \left(\frac{\partial G}{\partial P} \right)_{T,N} \implies \left(\frac{\partial S}{\partial P} \right)_T = - \left(\frac{\partial V}{\partial T} \right)_P.$$

These relations help us to obtain quantities in experiments when the quantity of interest is hard to determine. For example, $(\partial S/\partial V)_T$ is not easy to measure because measurements on entropy is needed, but its equivalent $(\partial P/\partial T)_V$ can tell you the answer. Note: it is in general not possible to determine both quantities in a *single* experiment, it is just the “numbers” that are equal. For example, to determine $(\partial S/\partial V)_T$ you need to vary the volume, but $(\partial P/\partial T)_V$ says that the volume is fixed. At least two experiments are needed.

5.2 Free Energy as a Force toward Equilibrium

The second law states that the entropy tends to increase for an isolated system. Then what if the system is not isolated, but is connected to and in thermal equilibrium with the environment—a large reservoir with unlimited energy and constant temperature? The total entropy of the system and the environment tends to increase. Consider a small change of the total entropy,

$$dS_{\text{total}} = dS + dS_R,$$

where S is the entropy of the system and S_R is the entropy of the environment. Using the thermodynamic identity (3.15), dS can be written as

$$dS = \frac{1}{T} dU + \frac{P}{T} dV - \frac{\mu}{T} dN.$$

Assuming V and N (and of course T) to be constant, then $dS_R = dU_R/T_R$, so

$$dS_{\text{total}} = dS + \frac{1}{T_R} dU_R.$$

The temperature of the reservoir is the same as the temperature of the system, and $dU_R = -dU$. Then

$$dS_{\text{total}} = dS - \frac{1}{T} dU = -\frac{1}{T} (dU - T dS) = -\frac{1}{T} dF.$$

In conclusion, under certain conditions (fixed T , V , and N), a decrease in the Helmholtz free energy of the system will increase the entropy of the universe. In other words, we can forget about the environment and just say that a system tends to minimize its Helmholtz free energy according to the second law. We can keep pressure constant and let the volume vary. By the same reasoning

$$dS_{\text{total}} = dS - \frac{1}{T} dU - \frac{P}{T} dV = -\frac{1}{T} (dU - T dS + P dV) = -\frac{1}{T} dG.$$

The Gibbs free energy of a system tends to decrease in order to satisfy the second law.

To sum up, whether the system is isolated or not, the second law requires that

- At constant energy and volume, S tends to increase.
- At constant temperature and volume, F tends to decrease.

- At constant temperature and pressure, G tends to decrease.

when we assume that no particle is leaving or entering the system. Please at least memorize the statement about G above. This will be the most useful one in the following sections in this chapter.

5.2.1 Extensive and Intensive Quantities

We now have many physical quantities: $U, V, N, S, T, P, \mu, H, F, G, \dots$. They can be classified into **extensive** and **intensive** quantities:

- Extensive: $V, N, S, U, H, F, G, m, \dots$
- Intensive: T, P, μ, ρ, \dots

An extensive quantity doubles when you make a copy of it. For example, when you bring two box of volume V together, their total volume is $2V$. If there are N particles in them, you will have $2N$ particles in total. An intensive quantity is invariant under the change in amount of stuff. A good example is the temperature: when you bring two system of temperature T in thermal contact, their temperature remains T .

Extensive quantities (E) and intensive quantities (I) obey the following rules:

$$E \times I = E, \quad E/E = I, \quad E + E = E, \quad I + I = I, \quad E \times E = \text{neither},$$

and $E + I$ is not allowed, even they have the same units (e.g. $U + \mu$ makes no sense). There certainly are other quantities that fall in neither categories, such as the multiplicity $\Omega = e^{S/k}$, which is an exponential of an extensive quantity.

5.2.2 Gibbs Free Energy and Chemical Potential

Recall that there is a partial-derivative relation for the Gibbs free energy according to (5.2),

$$\mu = \left(\frac{\partial G}{\partial N} \right)_{T,P}.$$

It is saying that adding a particle to a system, when keeping T and P fixed, increases the Gibbs free energy of the system by μ . You may think that μ may change when more particles are added, but it will not happen. This is because G is an extensive quantity and it must grow in proportion to the number of particles, or simply,

$$G = \sum_i N_i \mu_i \tag{5.4}$$

where i indicates the species of the particle/molecule. Such argument does not apply to the Helmholtz free energy, for

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

This time we are fixing T and V . Adding more and more particle does change μ because the system becomes more dense, while F is still an extensive quantity. You may observe a crucial difference in the two scenarios: we fix T and P (which are both intensive quantities) for the Gibbs free energy, but we fix T and V (which is extensive) for the Helmholtz energy. When intensive quantities are fixed, all extensive quantities grow in proportional to N . However, when extensive quantities are fixed while they *should* change with the growing amount of stuff, the proportionality gets messed up.

We may use (5.4) to do something. Consider a fixed amount of gas of the same species at a fixed temperature,

$$\frac{\partial \mu}{\partial P} = \frac{\partial}{\partial P} \left(\frac{G}{N} \right) = \frac{1}{N} \frac{\partial G}{\partial P} = \frac{V}{N}.$$

where we have used $(\partial G / \partial P)_{T,N} = V$. By the ideal gas law, this quantity is just kT/P ,

$$\frac{\partial \mu}{\partial P} = \frac{kT}{P}.$$

Multiplying both sides by dP and integrating gives

$$\mu(T, P) - \mu(T, P^\circ) = kT \ln \left(\frac{P}{P^\circ} \right),$$

where P° is some reference pressure (usually 1 atm). The chemical potential at this reference pressure is denoted as μ° , and thus

$$\mu(T, P) = \mu^\circ(T) + kT \ln \left(\frac{P}{P^\circ} \right). \quad (5.5)$$

For a mixture of ideal gases, (5.5) is to be calculated separately for each species and $P \rightarrow P_i$ would be the partial pressure,

$$\mu_i(T, P) = \mu_i^\circ(T) + kT \ln \left(\frac{P_i}{P^\circ} \right). \quad (5.6)$$

5.3 Phase Transformations of Pure Substances

Different forms of the substances, solid, liquid, and gas, are called **phases**. A **phase transformation** is a discontinuous change in the properties of a substance, as its environment is changed only infinitesimally. Both pressure and temperature may affect the phase of a substance. The equilibrium phases as a function of temperature and pressure is called a **phase diagram**. The phase diagram of water (H_2O) is shown in Figure (5.1). The lines on a phase diagram represent conditions under which two different phases can coexist in equilibrium.

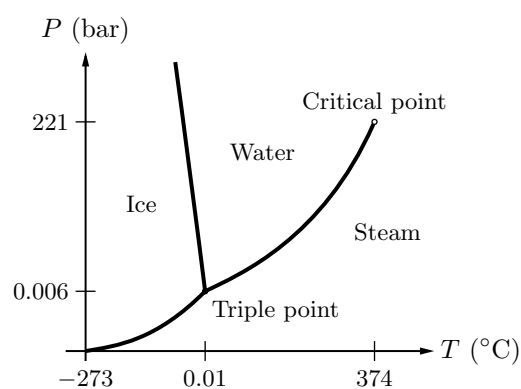


Figure 5.1: Phase diagram for H_2O .

The pressure at which a gas can coexist with its solid or liquid phase is called the **vapor pressure**. At the **triple point**, three phases can coexist in equilibrium (0.01°C and 0.006 bar for water). The liquid-gas equilibrium line terminates at a point called the **critical point**. Beyond (and at) this point, liquid phase and gas phase are nearly similar and the change from liquid to gas is continuous.

5.3.1 Diamonds and Graphite

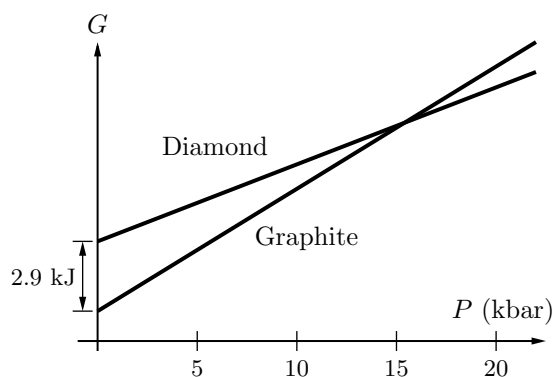


Figure 5.2: Molar Gibbs free energies of diamond and graphite as functions of pressure, at room temperature, neglecting the changes in volume.

Figure 5.2 shows a G - P graph of diamond and graphite. Similarly, the temperature dependence is

$$\left(\frac{\partial G}{\partial T} \right)_{P,N} = -S.$$

Elemental carbon has two common phases, namely diamond and graphite. At room temperature and pressure, graphite is more stable than diamond, so diamonds will spontaneously convert into graphite, but at an extremely slow rate. At high temperatures, this conversion is rapid. The reason for this is that the Gibbs free energy difference of graphite and diamond is about 2900 J/mol. At higher pressure the Gibbs free energy changes:

$$\left(\frac{\partial G}{\partial P} \right)_{T,N} = V.$$

A mole of graphite possess more volume than diamond, so its Gibbs free energy will grow faster than that of the diamond (and often the volume of these two are taken to be constant). Thus, at high pressure, diamond will be more stable than graphite.

The Gibbs free energy decreases more rapidly for graphite because it has more entropy, so raising the temperature makes graphite more stable than diamond.

5.3.2 The Clausius-Clapeyron Relation

We are interested in the function form of the equilibrium lines on the phase diagram. For convenience, consider the boundary between liquid and gas, which can be further generalized to other phase boundary. Let G_ℓ and G_g be the Gibbs free energies for liquid and solid, respectively. On the boundary,

$$G_\ell = G_g.$$

Now consider an increase of temperature by dT and pressure by dP so that the two phase are still in equilibrium. Then

$$dG_\ell = dG_g.$$

Using the thermodynamic identity (5.2) for G ,

$$-S_\ell dT + V_\ell dP = -S_g dT + V_g dP \implies \frac{dP}{dT} = \frac{S_g - S_\ell}{V_g - V_\ell}.$$

We often write $S_g - S_\ell = L/T$ where L is the latent heat (specific latent heat times mass) for phase transformation. Then we arrive at the **Clausius-Clapeyron relation**,

$$\frac{dP}{dT} = \frac{L}{T\Delta V} \quad (5.7)$$

Not just for liquid and gas, this can be used to determine any boundary lines on a phase diagram. Take water as an example. Water has a unusual negative slope of the equilibrium line between solid and liquid phase. This is because liquid water occupies less volume than ice ($\Delta V < 0$). Since the latent heat and the temperature are both positive, $dP/dT < 0$ for the boundary line of water and ice.

The Clausius-Clapeyron relation is a differential equation that can tell the entire phase-boundary curves if both $L(T)$ and $\Delta V(T)$ are known. For a reasonably small sections of a curve, and if one of the phases is gas, we can take L to be constant and ΔV just the volume of the gas. Then we will arrive at the vapor pressure equation using the ideal gas law $\Delta V \approx NkT/P$,

$$P = Ae^{-L/RT}, \quad (5.8)$$

where A is some constant.

5.3.3 The van der Waals Model

The **equation of state** of an ideal gas is the ideal gas law: $PV = NkT$. It does not say anything about phase transformation, as we can have any pressure or temperature. For a liquid-gas system, the **van der Waals equation** is a classic model that predicts phase transformations,

$$\left(P + \frac{aN^2}{V^2}\right)(V - Nb) = NkT. \quad (5.9)$$

This has the same structure as the ideal gas law, but with some correction terms. Each corrected term is associated with a modified assumption of an ideal gas.

Let us first look at the second modified term. It is just saying that a fluid cannot have a zero volume because actual gas molecules are *not* point particles (an ideal gas assume they are). Hence b is the minimum volume occupied by a molecule. The first modified term aN^2/V^2 represents the attraction force between molecules (while the ideal gas molecules does not have mutual attraction). An attraction force is associated with a negative potential energy. For a fixed volume, if we double the molecules inside, the potential energy of one molecule due to other molecules also doubles. This means the potential energy of *one* molecule is proportional to the number density N/V . There are in total N particles, so the total potential energy is proportional to N^2/V ,

$$\text{total potential energy} = -\frac{aN^2}{V},$$

where a is some positive constant. By the thermodynamic identity (3.12), $P = -(\partial U/\partial V)_S$, the pressure contribution from the

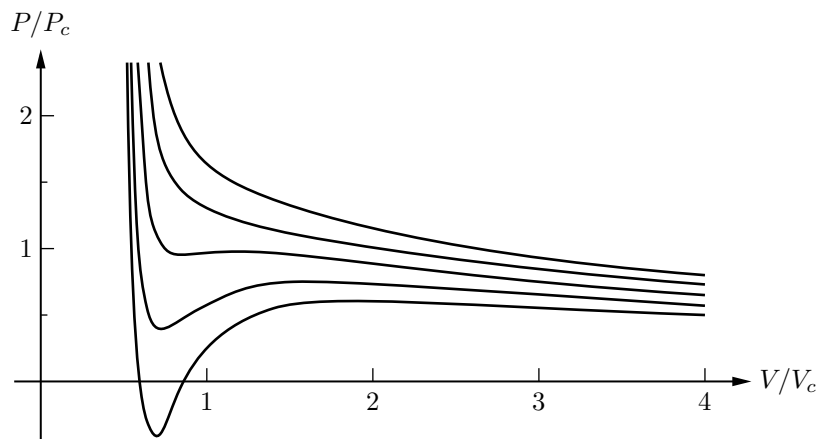


Figure 5.3: Isotherms for a van der Waals fluid. From bottom to top, the lines are for 0.8, 0.9, 1.0, 1.1, and 1.2 times T_c , the temperature at the critical point. The axes are labeled in units of the pressure and volume at the critical point; in these units the minimum volume (Nb) is $1/3$.

potential energy is

$$P_{\text{potential}} = -\frac{d}{dV} \left(-\frac{aN^2}{V} \right) = -\frac{aN^2}{V^2}.$$

Note that van der Waal's model is not an exact model. The constants a and b depends on the property of a substance and sometimes on different conditions. The consequence of the van der Waals model arises clearly on a P - V graph showing isotherms (see Figure 5.3).

When the temperature is low enough, P/P_c can be negative and something interesting might be going on. The region where pressure has a positive first derivative ($dP/dV > 0$) indicates that the system is unstable. Imagine that you push the piston on the gas by a little, the volume decreases and the pressure also decreases. The volume gets smaller and smaller until a collapse. How do we interpret these instability points? Recall that for a system in contact with the environment with constant temperature and pressure, the Gibbs free energy is to be minimized. For fixed N, T ,

$$dG = -S dT + V dP + \mu dN = V dP,$$

Dividing both sides by dV gives,

$$\left(\frac{\partial G}{\partial V} \right)_{N,T} = V \left(\frac{\partial P}{\partial V} \right)_{N,T} = -\frac{NkTV}{(V - Nb)^2} + \frac{2aN^2}{V^2}.$$

Integrating with respect to volume,

$$G = -NkT \ln(V - Nb) + \frac{(NkT)Nb}{V - Nb} - \frac{2aN^2}{V} + C(T),$$

where $C(T)$ is some integration constant, different for different temperatures. We will plot G vs. P , each parametrized by V . An example is shown in Figure 5.4 for a specific temperature.

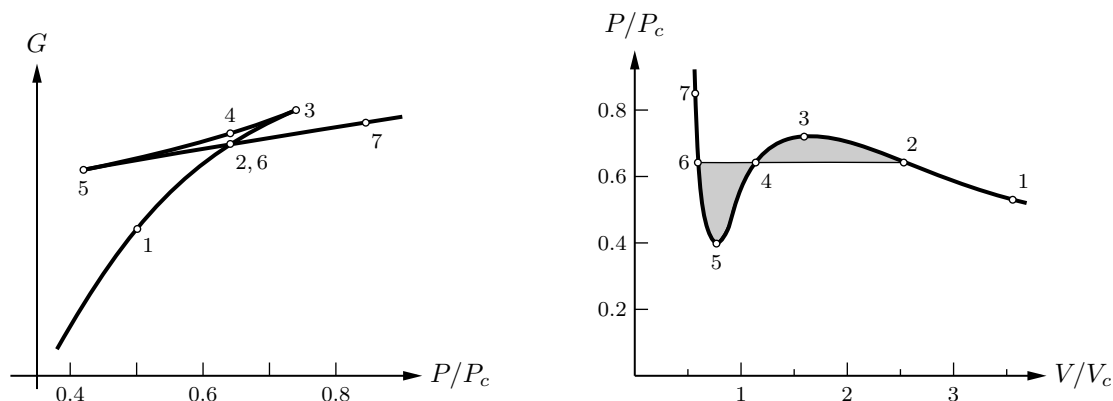


Figure 5.4: Gibbs free energy as a function of pressure for a van der Waals fluid at $T = 0.9 T_c$. The corresponding isotherm is shown at right. States in the range 2-3-4-5-6 are unstable.

Let us now analyze the graphs. Suppose the state starts out from point 1 and then reaches point 2. Will the state proceed to point 3 as van der Waals predicted? The answer is no, because point 3 is an unstable point for the Gibbs free energy. You can find a point with a lower G at the same pressure on that curve (somewhere between point 6-7). Similarly, point 4 and 5 are also unstable, and in fact all points around the loop 2-3-4-5-6 are unstable. The conclusion is: to minimize the Gibbs free energy along the curve, the state stays at point 2 and directly goes to point 6 on the G - P diagram. On the P - V diagram, starting from point 2, the state proceeds along the constant pressure line but with a gradually decreasing volume—a phase transformation. Once the state reaches point 6, the pressure rises abruptly as the volume increase, which indicates that liquid is highly incompressible. Between 2 and 6, the system is consist of a mixture of liquid and gas. In other words, pure liquid or gas is unstable at points between 2 and 6; a mixture state minimizes G .

Maxwell Construction

The constant pressure at phase transformation can be determined by the graph of G , but we can also get it form the P - V graph. The idea is that the two shaded areas are the same. We know that the change in Gibbs free energy around the loop (2-3-4-5-6) is zero because the start point and the end point are identical, so

$$0 = \int_{\text{loop}} dG = \int_{\text{loop}} \left(\frac{\partial G}{\partial P} \right)_T dP = \int_{\text{loop}} V dP.$$

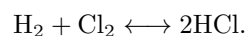
If you look sideways at the P - V graph, you will find this condition means the two shaded areas are equal. Drawing the straight pressure line to obtain two such equal areas is called the **Maxwell construction**. This pressure is called the **vapor pressure** where phase transformation from gas to liquid (or the converse) takes place.

From Figure 5.3, we see that not all isotherms have a vapor pressure line. For high temperature, the van der Waals gas looks just like an ideal gas, with no phase transformation. There is an isotherm that set the boundary temperature of this phenomena where both dP/dV and d^2P/dV^2 vanish. This is called the **critical temperature**, at which the volume is called the **critical volume** and the pressure is called the **critical pressure**. For a van der Waals gas, those critical parameters are

$$V_c = 3Nb, \quad P_c = \frac{1}{27} \frac{a}{b^2}, \quad \text{and} \quad kT_c = \frac{8}{27} \frac{a}{b}.$$

5.4 Chemical Equilibrium

Consider the following chemical reaction.



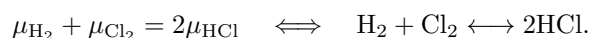
The reaction can go both sides. There exists an equilibrium state where the rates of the forward and backward reaction are equal. Most chemical reactions occur at fixed temperature and pressure, so the Gibbs free energy of the system is to be minimized as usual. By thermodynamic identity (5.2) for G ,

$$dG = 0 = \sum_i \mu_i dN_i \quad [\text{at equilibrium}],$$

where i indicates the species of interest. Note that dN_i 's are not independent. For $dN_{\text{H}_2} = 1$, there must be an associate $dN_{\text{Cl}_2} = 1$ and $dN_{\text{HCl}} = -2$ according to the chemical reaction formula. Thus, the equilibrium condition is

$$\sum_i \nu_i \mu_i = 0, \tag{5.10}$$

where ν_i is some stoichiometric coefficients (e.g. $\nu_{\text{H}_2} = \nu_{\text{Cl}_2} = 1$, $\nu_{\text{HCl}} = -2$ in the example). Simply, the equilibrium condition is just another way to write the chemical reaction,



Recall that for an ideal gas, the chemical potential and pressure are related by (5.6),

$$\mu(T, P) = \mu^\circ(T) + kT \ln \left(\frac{P}{P^\circ} \right),$$

where μ° is some standard chemical potential (which you can look up) at standard pressure. Then the equilibrium condition reads

$$\mu_{\text{H}_2}^\circ + kT \ln \left(\frac{P_{\text{H}_2}}{P^\circ} \right) + \mu_{\text{Cl}_2}^\circ + kT \ln \left(\frac{P_{\text{Cl}_2}}{P^\circ} \right) - 2\mu_{\text{HCl}}^\circ - 2kT \ln \left(\frac{P_{\text{HCl}}}{P^\circ} \right) = 0.$$

Rearranging some terms:

$$kT \ln \left(\frac{P_{\text{H}_2} P_{\text{Cl}_2}}{P_{\text{HCl}}^2} \right) = 2\mu_{\text{HCl}}^\circ - \mu_{\text{H}_2}^\circ - \mu_{\text{Cl}_2}^\circ,$$

and multiplying by Avogadro's number N_A ,

$$RT \ln \left(\frac{P_{\text{H}_2} P_{\text{Cl}_2}}{P_{\text{HCl}}^2} \right) = \Delta G^\circ.$$

The RHS is the standard Gibbs free energy ΔG° for a reaction. It is the change in Gibbs free energy of one mole of forward reaction at 1 bar. The above equation is equivalent to

$$\frac{P_{\text{HCl}}^2}{P_{\text{H}_2} P_{\text{Cl}_2}} = e^{-\Delta G^\circ / RT} \equiv K. \quad (5.11)$$

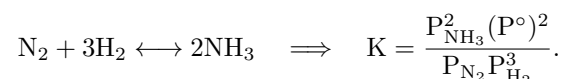
Here K is called the **equilibrium constant**. Equation (5.11) is for this specific reaction. The general equilibrium constant for any reaction (in gaseous state) with ν_i 's is

$$K = \prod_i \left(\frac{P^\circ}{P_i} \right)^{\nu_i}. \quad (5.12)$$

This equation is called the **law of mass action**. Let us see some qualitative results from this equation with the example of HCl. If

$$\frac{n_{\text{HCl}}^2}{n_{\text{H}_2} n_{\text{Cl}_2}} < K(T),$$

the reaction will run forward until the inequality becomes an equality. Similarly, if $n_{\text{HCl}}^2 / n_{\text{H}_2} n_{\text{Cl}_2} > K(T)$, the reaction will run in reverse. Another example is the nitrogen fixation process



If you double the partial pressure of H_2 and N_2 , then the partial pressure must quadruple to maintain equilibrium. This means increasing the total pressure of the system favors the production of more ammonia (and this is what we usually want). These response to disturbance of equilibrium is known as **Le Chatelier's principle**: When you disturb a system in equilibrium, it will respond in a way that partially offsets the disturbance.

Part II

Statistical Mechanics

6 BOLTZMANN STATISTICS

In chapter 4 and 5, we discussed some applications of thermodynamics using theoretical results derived in chapter 1-3. For these applications, experimental measurements are invoked (e.g. we need to look up tables for Gibbs free energy, enthalpy, etc.) It will be better to predict all these thermodynamic quantities from first principles. Until now, the models that we worked out from first principles only include: the two-state paramagnet, the Einstein solid, and the monatomic ideal gas. We wrote down the multiplicity Ω explicitly, then calculate the entropy, and then thermodynamic quantities like heat capacities. However, it is extremely difficult to directly get the multiplicity for more complicated models. To get quantities of interest, other theoretical tools are required, as will be discussed in this chapter.

6.1 The Boltzmann Factor

This section is about the *most powerful* tool in all of statistical mechanics. We want to know the probability of finding a system in any particular microstate, when that system is in thermal equilibrium with a “reservoir” at a specific temperature.

6.1.1 Boltzmann Distribution

Consider a single atom as the system. The microstate of such a system is then the energy levels of the atom. We will neglect spin for this moment. There are often more than one independent state for an energy level, which is called **degenerate**. For example, a hydrogen atom can have one ground state, but four independent first-excitation states, nine independent second-excitation states, and so on. Each independent state is a separate microstate.

Now, the atom is not isolated, but is exchanging energy with atoms from a large reservoir of temperature T . Some microstates will be more likely than others, depending on their energies. The key is that though the atom is not an isolated system, the atom with the reservoir is. Then according to fundamental assumption of statistical mechanics: it is equally likely to find this combined system in any of its accessible microstates. Suppose the total energy of the combined system is fixed. For a low energy atom, there will be more energy for the reservoir, in general larger multiplicity, and vice versa.

Consider two states s_1 and s_2 , with energies $E(s_1)$ and $E(s_2)$, probabilities $\mathcal{P}(s_1)$ and $\mathcal{P}(s_2)$, respectively. Let $\Omega_R(s_1)$ be the multiplicity of the reservoir when the atom is in state s_1 , and $\Omega_R(s_2)$ for state s_2 similarly. We want to find $\mathcal{P}(s_1)$ and $\mathcal{P}(s_2)$. First, we can immediately conclude that the ratio of probabilities for the two states is

$$\frac{\mathcal{P}(s_2)}{\mathcal{P}(s_1)} = \frac{\Omega_R(s_2)}{\Omega_R(s_1)}$$

because all microstates of the combined system are equally probable (and the multiplicity of the atom is just 1). In terms of entropy $S = k \ln \Omega$,

$$\frac{\mathcal{P}(s_2)}{\mathcal{P}(s_1)} = \frac{e^{S_R(s_2)/k}}{e^{S_R(s_1)/k}} = e^{[S_R(s_2) - S_R(s_1)]/k}.$$

By the thermodynamic identity (3.15):

$$dS_R = \frac{1}{T}(dU_R + P dV_R - \mu dN_R).$$

The $P dV_R$ term is often nonzero, but is negligible relative to dU_R . The dN_R term is literally zero because the reservoir is not gaining particles in this case (and in many other cases in this chapter). Thus, we can simply say

$$S_R(s_2) - S_R(s_1) = \frac{1}{T}[U_R(s_2) - U_R(s_1)] = -\frac{1}{T}[E(s_2) - E(s_1)],$$

and so

$$\frac{\mathcal{P}(s_2)}{\mathcal{P}(s_1)} = e^{-[E(s_2) - E(s_1)]/kT} = \frac{e^{-E(s_2)/kT}}{e^{-E(s_1)/kT}}.$$

Each exponential factor is called a **Boltzmann factor**:

$$\text{Boltzmann factor} = e^{-E(s)/kT}. \quad (6.1)$$

The Boltzmann factor is not the probability, but fortunately we can convert it into one. Note that

$$\frac{\mathcal{P}(s_2)}{e^{-E(s_2)/kT}} = \frac{\mathcal{P}(s_1)}{e^{-E(s_1)/kT}}.$$

Both sides need to be a constant, because the LHS does not depend on s_1 , then neither should the RHS. Likewise, both sides should not depend on s_2 . Each side must be a constant and it works for any state. We will define this constant to be $1/Z$. This converts a Boltzmann factor into a probability:

$$\mathcal{P}(s) = \frac{1}{Z} e^{-E(s)/kT} \quad (6.2)$$

This probability function is called the **Boltzmann distribution** or **canonical distribution**. It is the most useful formula in all of statistical mechanics. A system obeying Boltzmann statistics is called a **canonical ensemble**. A canonical ensemble has fixed particle number, volume, and temperature (instead of energy).

If you want the *total* probability ratio, don't forget to take into account the degeneracy:

$$\frac{N(s_2)}{N(s_1)} = \frac{n_2 \mathcal{P}(s_2)}{n_1 \mathcal{P}(s_1)}$$

where n_2 and n_1 are numbers of degenerate states for energy $E(s_2)$ and $E(s_1)$, respectively.

6.1.2 The Partition Function

To calculate Z , we use the normalization condition:

$$1 = \sum_s \mathcal{P}(s) = \sum_s \frac{1}{Z} e^{-E(s)/kT} = \frac{1}{Z} \sum_s e^{-E(s)/kT}.$$

Solving for Z ,

$$Z = \sum_s e^{-E(s)/kT}. \quad (6.3)$$

The quantity Z is called the **partition function**. It is a “constant” that depends on temperature, not on a state s . It can be an infinite series with some complicated formula of $E(s)$. Numerically, energies much greater than kT can be neglected, and you may just sum up the first several terms. For specific spacing among energy levels, Z depends on the ground state energy E_0 : $Z \rightarrow Z e^{E_0/kT}$. But this factor is unimportant, as it cancels when computing probabilities:

$$P(s) = \frac{1}{Z e^{E_0/kT}} e^{-[E(s)+E_0]/kT} = \frac{1}{Z} e^{-E(s)/kT}.$$

If there exist a ground state energy $E(0) = 0$, then at absolute zero, $Z = 1$.

6.2 Average Values

Now we get a way to express the probability. We can use it to calculate some average/expectation values. For a variable X with value $X(s)$ in state s , its average value is given by

$$\langle X \rangle \equiv \sum_s X(s) \mathcal{P}(s) = \frac{1}{Z} \sum_s X(s) e^{-\beta E(s)},$$

where $\beta = 1/kT$. For example, the first average value we calculate should be $\langle E \rangle$,

$$\langle E \rangle = \frac{\sum_s E(s) e^{-\beta E(s)}}{\sum_s e^{-\beta E(s)}}.$$

The numerator is very similar to the partition function, but with an extra factor of $E(s)$. One can cleverly take a derivative of Z with respect to β and find that

$$\langle E \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial}{\partial \beta} \ln Z = kT^2 \frac{\partial}{\partial T} \ln Z. \quad (6.4)$$

This is a very useful formula if you know the closed form of Z . The average value is nice, because once you have a large collection of identical, independent particles, you can compute their energy:

$$U = N\langle E \rangle.$$

The treatment of an atom and a reservoir is just a trick to find probabilities. We may see one particle as a system, treating the rest as the reservoir. The ultimate goal is to get the average value of the quantity of interest, and finally multiply by N . Again, we used the property of large numbers.

6.2.1 Paramagnetism

So much for the theory of Boltzmann statistics. Let us look at some examples. In section 3.3 about the Ising paramagnet, we derived its multiplicity explicitly, and then got the entropy, temperature, heat capacity, etc. We will see how Boltzmann statistics gives us thermodynamic quantities directly.

Recall that we define the energy of one dipole in an ideal two-state paramagnet to be

$$E_{\uparrow} = -\mu B \quad \text{and} \quad E_{\downarrow} = \mu B.$$

Take one dipole as the system and rest of the dipoles as the environment, we will know the probability of finding one of the two states:

$$\mathcal{P}_{\uparrow} = \frac{1}{Z} e^{\beta\mu B} \quad \text{and} \quad \mathcal{P}_{\downarrow} = \frac{1}{Z} e^{-\beta\mu B}, \quad \text{where} \quad Z = e^{\beta\mu B} + e^{-\beta\mu B} = 2 \cosh(\beta\mu B).$$

Using (6.4) (you can also do this explicitly by definition of average value), the average energy is simply

$$\langle E \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{1}{2 \cosh(\beta\mu B)} [2\mu B \sinh(\beta\mu B)] = -\mu B \tanh(\beta\mu B).$$

The total energy is consistent with Equation (3.7):

$$U = -N\mu B \tanh(\beta\mu B).$$

Another important quantity is the magnetization. First find the average value of the dipole moment,

$$\langle \mu \rangle = \mu \mathcal{P}_{\uparrow} + (-\mu) \mathcal{P}_{\downarrow} = \mu \left(\frac{e^{\beta\mu B} - e^{-\beta\mu B}}{e^{\beta\mu B} + e^{-\beta\mu B}} \right) = \mu \tanh(\beta\mu B).$$

The magnetization is also consistent with Equation (3.8):

$$M = N\langle \mu \rangle = N\mu \tanh(\beta\mu B).$$

6.2.2 Rotation and Vibration of Diatomic Molecules

Consider a diatomic molecule made up of atoms of different types (e.g. CO) for now. Its internal energy consists of binding energy, vibrational energy, rotational energy.

$$E_{\text{internal}} = E_{\text{bind}} + \left(n + \frac{1}{2} \right) \hbar\omega + j(j+1) \frac{\hbar^2}{2I}.$$

The first term is unimportant; the second term is the energy level of a quantum harmonic oscillator ($n = 0, 1, 2, 3, \dots$); the third term is the energy associated with angular momentum of the particle ($j = 0, 1, 2, 3, \dots$). We can divide the partition function into three pieces:

$$\begin{aligned} Z &= \sum_s \exp \left[-\frac{1}{kT} \left(E_{\text{gs}} + n\hbar\omega + j(j+1) \frac{\hbar^2}{2I} \right) \right] \\ &\rightarrow e^{-E_{\text{gs}}/kT} \cdot \sum_n e^{-n\hbar\omega/kT} \cdot \sum_j (2j+1) e^{-j(j+1)\hbar^2/2IkT} \\ &= Z_{\text{gs}} \cdot Z_{\text{vib}} \cdot Z_{\text{rot}}. \end{aligned}$$

Here the ground state energy of the harmonic oscillator ($\hbar\omega/2$) is absorbed into the binding energy, where we call it E_{gs} (ground-state energy). The $(2j+1)$ factor in Z_{rot} comes from the fact that the number of degenerate states for j th level is $2j+1$.

Rotational Energy

The partition function of rotational energy is

$$Z = \sum_{j=0}^{\infty} (2j+1) e^{-j(j+1)\epsilon/kT},$$

where $\epsilon = \hbar^2/2I$. This sum cannot be evaluated in a closed form, but we can do it numerically. At high temperature, $e^{-\epsilon/kT}$ decays slowly and there are more terms for Z , so the sum can be approximated as an integral. Figure 6.1 shows a bar graph of the partition function and a smooth fitting curve at $kT/\epsilon = 30$.

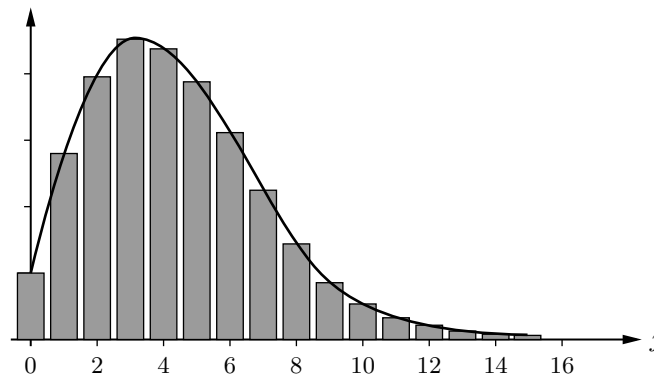


Figure 6.1: Bar-graph representation of the rotational partition function. At high temperatures the sum can be approximated as the area under a smooth curve.

The integral is very easy to evaluate by making substitutions

$$Z_{\text{rot}} \approx \int_0^{\infty} (2j+1) e^{-j(j+1)\epsilon/kT} dj = \frac{kT}{\epsilon} \int_0^{\infty} e^{-u} du = \frac{kT}{\epsilon},$$

where $u = j(j+1)\epsilon/kT$ and $du = (2j+1)\epsilon/kT$. Now we have a partition function in the high temperature limit, in terms of $\beta = 1/kT$. The average rotational energy is

$$\langle E_{\text{rot}} \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -(\beta\epsilon) \frac{\partial}{\partial \beta} \frac{1}{\beta\epsilon} = \frac{1}{\beta} = kT, \quad [kT \gg \epsilon].$$

This is what equipartition theorem predicts: diatomic molecule has two rotational degrees of freedom, so the average energy of each molecule is $2 \cdot kT/2 = kT$. This also explain why rotational degrees of freedom is available only at sufficient temperature (see Figure 1.2). At low temperature, these degrees of freedom “freeze out”. For molecules made of identical particles, rotating the molecule by 180° does nothing. The total states of the molecule is halved, so does the partition function (at high-temperature limit),

$$Z_{\text{rot}} \approx \frac{kT}{2\epsilon}, \quad [\text{identical atoms, } kT \gg \epsilon].$$

Note that this $1/2$ does not affect the average energy and the heat capacity as you re-calculate $\langle E_{\text{rot}} \rangle$.

Vibrational Energy

Using the same method as in rotational energy, the partition function for the vibrational energy can also be approximated as an integral at high temperature,

$$Z_{\text{vib}} = \sum_n e^{n\hbar\omega/kT} \approx \int_0^{\infty} e^{-n\hbar\omega/kT} dn = \frac{kT}{\hbar\omega}.$$

It takes the same form as the one of rotational energy, so the average energy is also

$$\langle E_{\text{vib}} \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -(\beta\hbar\omega) \frac{\partial}{\partial \beta} \frac{1}{\beta\hbar\omega} = kT.$$

Equipartition theorem also predicts this as in Figure 1.2: the vibrational energy contains kinetic and potential energy, both are quadratic degrees of freedom. Note that vibrational energy is available only at even higher temperature, because $\hbar\omega$ is usually greater than ϵ . Another approach to vibrational energy is available in Appendix A.2.1.

6.3 The Equipartition Theorem

Now with the tool of Boltzmann statistics, it is time to prove the equipartition theorem that has been used in so many chapters. The equipartition theorem applies to energy with quadratic degrees of freedom, so we should assume the energy to be of the form

$$E(q) = cq^2,$$

where c is some constant and q is any coordinate or momentum variable (e.g. x, p_x, L_x). Consider a system with one quadratic degree of freedom (which can be generalized to f quadratic degrees of freedom) q . Assume the system is in equilibrium with a reservoir at temperature T , and q is discrete but separated by small intervals δq . The partition function for the system is

$$Z = \sum_q e^{-\beta E(q)} = \sum_q e^{-\beta cq^2} = \frac{1}{\delta q} \sum_q e^{-\beta cq^2} \delta q.$$

Since δq is very small, the Riemann sum can be approximated as an integral:

$$Z \approx \frac{1}{\delta q} \int_{-\infty}^{+\infty} e^{-\beta cq^2} dq = \frac{1}{\delta q} \sqrt{\frac{\pi}{\beta c}} = C\beta^{-1/2},$$

where $C = \sqrt{\pi/c}/\delta q$. Using (6.4) to find the average value of energy:

$$\langle E \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{1}{C\beta^{-1/2}} \frac{\partial}{\partial \beta} C\beta^{-1/2} = -\frac{1}{\beta^{-1/2}} \left(-\frac{1}{2} \beta^{-3/2} \right) = \frac{1}{2\beta} = \frac{1}{2} kT.$$

This is the equipartition theorem. Note that our derivation relies on the fact that δq is very small. If δq , or the spacing between energy levels is comparable to kT , equipartition theorem fails. For example, the equipartition theorem is only true for an Einstein solid in the high-temperature limit, i.e. $kT \gg \hbar\omega$. In quantum statistics, the equipartition theorem is no longer a good approximation of energy.

You can use Boltzmann statistics to find other “equipartition theorems” for different kinds of degrees of freedom. For example, the “equipartition theorem for linear degrees of freedom” is $\langle E \rangle = kT$ (see Appendix A.2.2).

6.4 The Maxwell Speed Distribution

6.4.1 Speed Distribution

From the equipartition theorem, we already know that the root-mean-square speed of particles in an ideal gas is

$$v_{\text{rms}} = \sqrt{\frac{3kT}{m}}.$$

As its name suggests, this is just some average speed. The distribution of speeds is unknown from this expression. We will use Boltzmann statistics to derive the speed distribution, the probability density of some particle moving at a given speed, of an ideal gas.

The distribution function $\mathcal{D}(v)$, like any other probability density, is normalized. The probability of finding the particle’s speed between v_1 and v_2 is

$$\mathcal{P}(\text{between } v_1 \text{ and } v_2) = \int_{v_1}^{v_2} \mathcal{D}(v) dv.$$

The space is three dimensional, so for a given speed, there can be many possible velocity vectors. This means

$$\mathcal{D}(v) \propto \left(\begin{array}{c} \text{probability of a particle} \\ \text{having velocity } \mathbf{v} \end{array} \right) \times \left(\begin{array}{c} \text{number of vectors } \mathbf{v} \\ \text{corresponding to speed } v \end{array} \right).$$

We know that the velocity is related to kinetic energy by $E = \frac{1}{2}mv^2$. Hence the probability of having \mathbf{v} is proportional to the Boltzmann factor $e^{-mv^2/2kT}$. For the second factor, consider a three-dimensional velocity space. Any velocity vector \mathbf{v} corresponding to speed v lies on the surface of a sphere with radius v in this space. We see that a large v has more possible velocity vectors because

it corresponds to a larger sphere. Thus, the number of \mathbf{v} corresponding to speed v is proportional to $4\pi v^2$. Combining the two factors, we have

$$\mathcal{D}(v) = C \cdot 4\pi v^2 e^{-mv^2/2kT}.$$

To determine C , use the normalization condition for a distribution,

$$1 = \int_0^\infty \mathcal{D}(v) dv = 4\pi C \int_0^\infty v^2 e^{-mv^2/2kT} dv.$$

Making a change of variables $x = v\sqrt{m/2kT}$, this equation becomes

$$1 = 4\pi C \left(\frac{2kT}{m}\right)^{3/2} \int_0^\infty x^2 e^{-x^2} dx = 4\pi C \left(\frac{2kT}{m}\right)^{3/2} \frac{\sqrt{\pi}}{4}.$$

This gives $C = (m/2\pi kT)^{3/2}$. The final result for the distribution function is

$$\mathcal{D}(v) = \left(\frac{m}{2\pi kT}\right)^{3/2} 4\pi v^2 e^{-mv^2/2kT}. \quad (6.5)$$

This is known as the **Maxwell speed distribution** for an ideal gas.

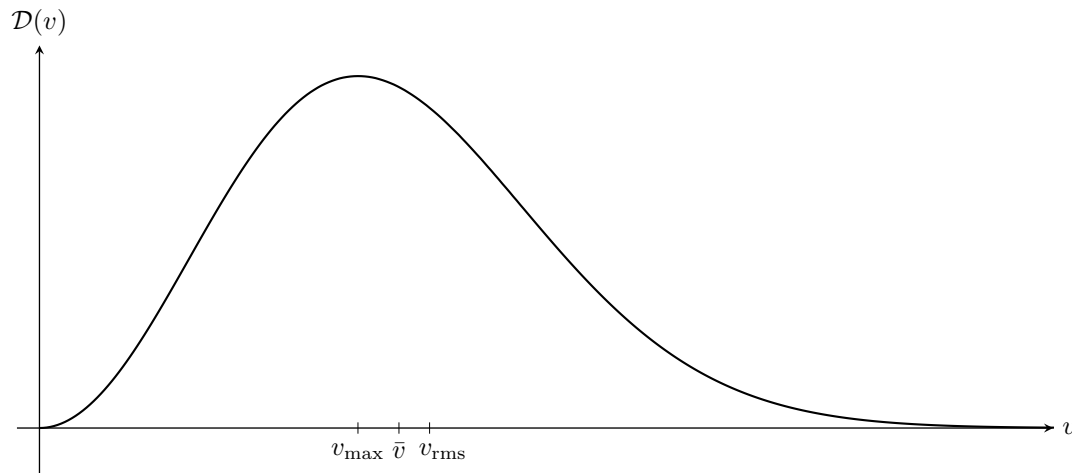


Figure 6.2: The Maxwell speed distribution falls off as $v \rightarrow 0$ and as $v \rightarrow \infty$. The average speed is slightly larger than the most likely speed, while the rms speed is a even more larger.

At small v , the probability looks like a parabola. At large v , the probability is exponentially decaying. There are two more speeds we can draw from this graph. One is the most probable speed v_{\max} . This is obtained by taking the derivative of the distribution:

$$\frac{d\mathcal{D}}{dv} \propto \frac{d}{dv} \left(v^2 e^{-mv^2/2kT} \right) = 2v e^{-mv^2/2kT} - v^2 \left(\frac{mv}{kT} \right) e^{-mv^2/2kT} = 2v \left(1 - \frac{mv^2}{2kT} \right) e^{-mv^2/2kT}.$$

Setting this equal to zero gives $v = 0$ or $v^2 = 2kT/m$. At $v = 0$, the distribution is at minimum, so the most probable speed is at $v_{\max} = \sqrt{2kT/m}$.

The average speed \bar{v} is determined by the following expression:

$$\bar{v} = \int_0^\infty v \mathcal{D}(v) dv = \left(\frac{m}{2\pi kT}\right)^{3/2} 4\pi \int_0^\infty v^3 e^{-mv^2/2kT} dv.$$

To do this integral analytically, you can first evaluate the integral of $ve^{-\lambda v^2}$, where $\lambda = -m/2kT$,

$$\int_0^\infty v e^{-\lambda v^2} dv = \frac{1}{2} \int_0^\infty e^{-\lambda v^2} d(v^2) = -\frac{1}{2\lambda} e^{-\lambda v^2} \Big|_0^\infty = \frac{1}{2\lambda}.$$

Taking the derivative with respect to λ gives the desired integral

$$\int_0^\infty v^3 e^{-\lambda v^2} dv = -\frac{\partial}{\partial \lambda} \int_0^\infty v e^{-\lambda v^2} dv = -\frac{\partial}{\partial \lambda} \left(\frac{1}{2\lambda} \right) = \frac{1}{2\lambda^2}.$$

Plugging in $\lambda = -m/2kT$, average speed is

$$\bar{v} = \left(\frac{m}{2\pi kT}\right)^{3/2} 4\pi \int_0^\infty v^3 e^{-mv^2/2kT} dv = \left(\frac{m}{2\pi kT}\right)^{3/2} 4\pi \cdot \frac{1}{2} \left(\frac{2kT}{m}\right)^2 = \sqrt{\frac{8kT}{\pi m}}.$$

Note that $v_{\max} < \bar{v} < v_{\text{rms}}$. You may wonder why the average speed is not v_{rms} . This is because the root-mean-square speed is defined (or related) to the average of speed squared. To see this, let's calculate the average of v^2 and see what we get:

$$\begin{aligned} \langle v^2 \rangle &= \int_0^\infty v^2 \mathcal{D}(v) dv = \frac{4}{\sqrt{\pi}} \left(\frac{m}{2kT}\right)^{3/2} \int_0^\infty v^4 e^{-mv^2/2kT} dv \\ &= \frac{4}{\sqrt{\pi}} \left(\frac{m}{2kT}\right)^{3/2} \left(\frac{2kT}{m}\right)^{5/2} \int_0^\infty x^4 e^{-x^2} dx = \frac{8kT}{\sqrt{\pi m}} \int_0^\infty x^4 e^{-x^2} dx \\ &= \frac{8kT}{\sqrt{\pi m}} \cdot \frac{3\sqrt{\pi}}{8} = \frac{3kT}{m}. \end{aligned}$$

This is consistent with the equipartition theorem.

6.4.2 Energy Distribution

It is useful to convert the speed distribution to energy distribution, in case we want to know the most probable energy. The energy and speed are related by $E = \frac{1}{2}mv^2$, or $v = \sqrt{2E/m}$. In deriving speed distribution, we have

$$\mathcal{D}(v) = C \cdot 4\pi v^2 e^{-mv^2/2kT}.$$

Substituting $v = \sqrt{2E/m}$ and $C = (m/2\pi kT)^{3/2}$ into the normalization condition,

$$1 = 4\pi C \int_0^\infty v^2 e^{-mv^2/2kT} dV = 4\pi C \int_0^\infty \left(\frac{2E}{m}\right) \sqrt{\frac{1}{2mE}} e^{-E/kT} dE = 2 \left(\frac{1}{kT}\right)^{3/2} \int_0^\infty \sqrt{\frac{E}{\pi}} e^{-E/kT} dE.$$

Essentially, this is saying that the energy distribution is

$$\mathcal{D}(E) = 2\sqrt{\frac{E}{\pi}} \left(\frac{1}{kT}\right)^{3/2} e^{-E/kT}. \quad (6.6)$$

6.5 Partition Functions and Free Energy

The multiplicity $\Omega(U)$ is the most fundamental quantity of an isolated system with fixed energy U . For a system with fixed temperature T , the partition function $Z(T)$ is analogous to the multiplicity. We might expect that $\ln Z$ tends to increase. Under constant temperature, we also know that the Helmholtz free energy F tends to decrease, or $-F$ tends to increase. With the right dimensions, we can take a bold guess:

$$F = -kT \ln Z \quad (6.7)$$

Proof. Recall the definition of Helmholtz free energy and its first derivative:

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

Combining these two gives a differential equation for $F(T)$,

$$\frac{\partial F}{\partial T} = \frac{F - U}{T}.$$

We will find that the quantity $-kT \ln Z$ obeys the same differential equation. Let $\tilde{F} = -kT \ln Z$. Take the derivative with respect to temperature,

$$\frac{\partial \tilde{F}}{\partial T} = \frac{\partial}{\partial T}(-kT \ln Z) = -k \ln Z - kT \frac{\partial}{\partial T} \ln Z.$$

The second term looks familiar—it is proportional to the average energy according to (6.4). Replacing $\langle E \rangle$ with U (because we are

considering Z of a large system),

$$\frac{\partial \tilde{F}}{\partial T} = -k \ln Z - \frac{U}{T} = \frac{\tilde{F}}{T} - \frac{U}{T}.$$

A first order differential equation needs one initial condition to specify the solution. In this case, if $F(T \rightarrow 0) = \tilde{F}(T \rightarrow 0)$, we can prove that $F = \tilde{F}$ as they obey the same differential equation. As $T \rightarrow 0$, $F = U$ and this U is the lowest possible energy U_0 because all other boltzmann factors $e^{-U(s)/kT}$ are infinitely suppressed comparing to the ground state. Similarly, as $T \rightarrow 0$, $Z = e^{-U_0/kT}$, and

$$\tilde{F}(0) = -kT \ln Z(0) = U_0 = F(0).$$

□

The relation (6.7) is useful because once we get the partition function, other thermodynamic quantities are known,

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

6.6 Partition Functions for Composite Systems

Consider a system of two particles that do not have mutual interactions (like an ideal gas). The total energy of the system is $E_1 + E_2$, so the partition function reads

$$Z_{\text{total}} = \sum_s e^{-\beta[E_1(s) + E_2(s)]} = \sum_s e^{-\beta E_1(s)} e^{-\beta E_2(s)},$$

where s is all possible state of the *system*. If the two particles are distinguishable, then

$$Z_{\text{total}} = \sum_{s_1} \sum_{s_2} e^{-\beta E_1(s_1)} e^{-\beta E_2(s_2)} = \sum_{s_1} e^{-\beta E_1(s_1)} \sum_{s_2} e^{-\beta E_2(s_2)} = Z_1 Z_2.$$

Here Z_1 is the partition function of particle 1 alone, likewise for Z_2 . If the particles are *indistinguishable*, swapping the character of particle 1 and particle 2 does nothing to the system. This means we are counting the number of states nearly twice, so

$$Z_{\text{total}} = \frac{1}{2} Z_1 Z_2 = \frac{1}{2} Z_1^2 \quad [\text{noninteracting, indistinguishable particles}].$$

We replace Z_2 by Z_1 because they are identical particles. If $Z_1 \neq Z_2$ then we would be able to tell them apart, i.e. they become distinguishable. Note that this Z_{total} for indistinguishable particles are still not accurate. For two distinguishable particles, when $s_1 = s_2$, exchanging them still produce the same state. Then for indistinguishable particles, we are not overcounting these kind of states. But for large systems where the density is low enough (like an ideal gas) that the chances of both particles being in the same state are negligible, $Z_{\text{total}} = Z_1^2/2$ is quite accurate. We can generalize the partition function to a system with N particles:

$$Z_{\text{total}} = \begin{cases} Z_1 Z_2 Z_3 \cdots Z_N & [\text{noninteracting, distinguishable particles}], \\ \frac{1}{N!} Z_1^N & [\text{noninteracting, indistinguishable particles}]. \end{cases}$$

For a system with N noninteracting indistinguishable particles, we can calculate its Helmholtz free energy and the chemical potential.

$$F = -kT \ln Z = -kT [N \ln Z_1 - \ln N!] \approx -kT [N \ln Z_1 - N \ln N + N] = NkT \left[\ln \frac{Z_1}{N} + 1 \right].$$

The chemical potential is

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{T,V} = -kT \left[\ln \frac{Z_1}{N} + 1 \right] - NkT \frac{\partial}{\partial N} (-\ln N) = -kT \ln \frac{Z_1}{N}. \quad (6.8)$$

6.7 Ideal Gas Revisited

6.7.1 The Partition Function

An ideal gas containing N identical molecules with no interaction energy has the partition function of the form

$$Z = \frac{1}{N!} Z_1^N \quad (6.9)$$

The Boltzmann factor of a molecule can be divided into two parts:

$$e^{-E(s)/kT} = e^{-E_{\text{tr}}(s)/kT} e^{-E_{\text{int}}/kT},$$

where E_{tr} is the translational kinetic energy and E_{int} is its **internal energy** (rotational, vibrational, etc.) for state s . As in the last section, the sum in the partition function can be separated into two independent sums,

$$Z_1 = Z_{\text{tr}} Z_{\text{int}}, \quad \text{where} \quad Z_{\text{tr}} = \sum_{s_{\text{tr}}} e^{-E_{\text{tr}}/kT} \quad \text{and} \quad Z_{\text{int}} = \sum_{s_{\text{int}}} e^{-E_{\text{int}}/kT}.$$

The internal partition function are discussed in Section 6.2, so we should focus on the translational part. Suppose the molecules are confined in a three-dimensional box (fixed volume). We will first analyze the one-dimensional box of length L with one molecule in it. This is the famous “infinite square well” potential or “particle in a box” in quantum mechanics. The potential (energy) is zero inside the box, but infinite elsewhere. The wavefunction obtained by solving the Schrödinger equation with this potential has the form

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) \quad \text{with} \quad E_n = \frac{n^2 \pi^2 \hbar^2}{2mL^2} = \frac{n^2 \hbar^2}{8mL^2}.$$

These are stationary states with definite energies. The wavefunctions of infinite square well are standing waves with wavelength $\lambda_n = 2L/n$ as you can see from its functional form. With the energies, we can write the partition function:

$$Z_{1d} = \sum_n e^{-E_n/kT} = \sum_n e^{-\hbar^2 n^2 / 8mL^2 kT}.$$

As long as L and T are large, the energy levels are close so that the sum can be approximated as an integral,

$$Z_{1d} = \int_0^\infty e^{-n^2 \hbar^2 / 8mL^2 kT} dn = \frac{\sqrt{\pi}}{2} \sqrt{\frac{8mL^2 kT}{\hbar^2}} = \sqrt{\frac{2\pi m kT}{\hbar^2}} L \equiv \frac{L}{\ell_Q},$$

where

$$\ell_Q \equiv \frac{\hbar}{\sqrt{2\pi m kT}}$$

is called the **quantum length**. This is the de Broglie wavelength of a particle of mass m with kinetic energy kT , off by a factor of $\sqrt{\pi}$. For a nitrogen molecule at room temperature, $\ell_Q \approx 1.9 \times 10^{-11}$ m, so L/ℓ_Q is large for a practical box. Generalizing two three dimensional box, the total energy of a molecule is

$$E_{\text{tr}} = \frac{n_x^2 \hbar^2}{8mL_x^2} + \frac{n_y^2 \hbar^2}{8mL_y^2} + \frac{n_z^2 \hbar^2}{8mL_z^2}.$$

The partition function turns out to be

$$Z_{\text{tr}} = \sum_s e^{-E_{\text{tr}}/kT} = \frac{L_x}{\ell_Q} \frac{L_y}{\ell_Q} \frac{L_z}{\ell_Q} = \frac{V}{v_Q},$$

where V is the volume of the box and $v_Q = \ell_Q^3$ is the **quantum volume**. Together with the internal partition function, the total partition function of N molecules in a volume V is

$$Z = \frac{1}{N!} Z_1^N = \frac{1}{N!} \left(\frac{V Z_{\text{int}}}{v_Q} \right)^N, \quad \text{where} \quad Z_1 = \frac{V}{v_Q} Z_{\text{int}}. \quad (6.10)$$

For reference, after some Stirling’s approximation,

$$\ln Z = N[\ln V + \ln Z_{\text{int}} - \ln N - \ln v_Q + 1],$$

and the Helmholtz free energy is

$$\begin{aligned} F &= -kT \ln Z = -NkT[\ln V + \ln Z_{\text{int}} - \ln N - \ln v_Q + 1] \\ &= -NkT[\ln V - \ln N - \ln v_Q + 1] + F_{\text{int}}, \end{aligned}$$

where $F_{\text{int}} = -NkT \ln Z_{\text{int}}$ is the internal contribution to F .

6.7.2 Predictions

We have the partition function in a closed form. Let us calculate some thermodynamic quantities.

- Total average energy:

$$U = -\frac{\partial}{\partial \beta} \ln Z = -N \frac{\partial}{\partial \beta} \ln Z_{\text{int}} + N \frac{1}{v_Q} \frac{\partial v_Q}{\partial \beta} = N \langle E_{\text{int}} \rangle + N \cdot \frac{3}{2} \frac{1}{\beta} = U_{\text{int}} + \frac{3}{2} NkT.$$

The average translational kinetic energy is $U_{\text{tr}} = \frac{3}{2} NkT$ as expected.

- Pressure:

$$P = -\left(\frac{\partial F}{\partial V}\right)_{T,N} = \frac{NkT}{V}.$$

- Entropy: note that

$$v_Q = \left(\frac{h^2}{2\pi mkT}\right)^{3/2} \implies \ln v_Q = -\frac{3}{2} \ln T + \frac{3}{2} \ln \left(\frac{2\pi mk}{h^2}\right) = -\frac{3}{2} \ln T + \text{const.}$$

The entropy is

$$\begin{aligned} S &= -\left(\frac{\partial F}{\partial T}\right)_{V,N} = Nk \left[\ln \left(\frac{V}{Nv_Q} \right) + 1 \right] + (NkT) \left(\frac{3}{2} \frac{1}{T} \right) - \frac{\partial F_{\text{int}}}{\partial T} \\ &= Nk \left[\ln \left(\frac{V}{N} \left(\frac{2\pi mkT}{h^2} \right)^{3/2} \right) + \frac{5}{2} \right] + Nk \ln Z_{\text{int}}. \end{aligned}$$

By substituting $U_{\text{tr}} = \frac{3}{2} NkT$, we will recover the Sackur-Tetrode equation (2.10) for a monatomic ideal gas if there is no internal contributions.

- Chemical potential:

$$\begin{aligned} \mu &= \left(\frac{\partial F}{\partial N}\right)_{T,V} = -kT \left[\ln \left(\frac{V}{Nv_Q} \right) + 1 \right] + (NkT) \frac{1}{N} + \frac{\partial F_{\text{int}}}{\partial N} \\ &= -kT \ln \left[\frac{V}{N} \left(\frac{2\pi mkT}{h^2} \right)^{3/2} \right] - kT \ln Z_{\text{int}}. \\ &= -kT \ln \left(\frac{VZ_{\text{int}}}{Nv_Q} \right). \end{aligned}$$

For a monatomic ideal gas, there is no internal contribution. Equation (3.14) is recovered.

- Gibbs free energy: It is convenient to compute the Gibbs free energy of an ideal gas when we know its Helmholtz free energy. Just add $PV = NkT$ to F :

$$G = F + PV = -NkT \left[\ln \left(\frac{VZ_{\text{int}}}{Nv_Q} \right) + 1 \right] + NkT = -NkT \ln \left(\frac{VZ_{\text{int}}}{Nv_Q} \right) = N\mu.$$

7 QUANTUM STATISTICS

7.1 The Gibbs Factor

In Section 6.1, when deriving the Boltzmann factor, the small system is only allowed to exchange energy with the reservoir. What happens if the small system is able to exchange particles with the reservoir? As usual, we can write the ratio of probabilities for two states as

$$\frac{\mathcal{P}(s_2)}{\mathcal{P}(s_1)} = \frac{e^{S_R(s_2)/k}}{e^{S_R(s_1)/k}} = e^{[S_R(s_2) - S_R(s_1)]/k}.$$

Invoking the thermodynamic identity (3.15) to relate entropy with other quantities:

$$dS_R = \frac{1}{T}(dU_R + P dV_R - \mu dN_R).$$

This time, dN is no longer zero. Only $P dV$ is negligible, and the change in entropy becomes

$$S_R(s_2) - S_R(s_1) = -\frac{1}{T}[E(s_2) - E(s_1) - \mu N(s_2) + \mu N(s_1)].$$

Note that all quantities on the RHS don't have the subscript R . They refer to quantities of the small system. The probability ratio is now

$$\frac{\mathcal{P}(s_2)}{\mathcal{P}(s_1)} = \frac{e^{-[E(s_2) - \mu N(s_2)]/kT}}{e^{-[E(s_1) - \mu N(s_1)]/kT}}.$$

Each new exponential factor is called a **Gibbs factor**:

$$\text{Gibbs factor} = e^{-[E(s) - \mu N(s)]/kT}. \quad (7.1)$$

Like the Boltzmann factor, the Gibbs factor does not represent a probability. We will introduce the normalization constant $1/\mathcal{Z}$ and make the Gibbs factor a probability:

$$\mathcal{P}(s) = \frac{1}{\mathcal{Z}} e^{-[E(s) - \mu N(s)]/kT}. \quad (7.2)$$

where the quantity \mathcal{Z} , analogous to the partition function, is called the **grand partition function** or the **Gibbs sum**,

$$\mathcal{Z} = \sum_s e^{-[E(s) - \mu N(s)]/kT}. \quad (7.3)$$

If more than one type of particle is present in the system, then μdN becomes $\sum_i \mu_i dN_i$. The modified Gibbs factor is

$$\text{Gibbs factor} = e^{-[E(s) - \mu_A N_A(s) - \mu_B N_B(s) - \dots]/kT}.$$

A system with fixed temperature, chemical potential, and volume is called a **grand canonical ensemble**.

7.1.1 Average Values

Just like how we find the average energy by taking derivative of the partition function,

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

we can now compute the average particle in the system using the grand partition function:

$$\langle N \rangle = \sum_s N(s) \mathcal{P}(s) = \frac{\sum_s N(s) e^{-[E(s) - \mu N(s)]/kT}}{\sum_s e^{-[E(s) - \mu N(s)]/kT}}.$$

The numerator looks a lot like the grand partition function. Just take the derivative of it with respect to μ :

$$\langle N \rangle = \frac{kT}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial \mu} = kT \frac{\partial}{\partial \mu} \ln \mathcal{Z}. \quad (7.4)$$

7.1.2 The Grand Free Energy

From Boltzmann statistics, we know that the Helmholtz free energy is closely connected to the partition function ($F = -kT \ln Z$). Its counterpart in quantum statistics is the **grand free energy**:

$$\Phi \equiv U - TS - \mu N. \quad (7.5)$$

The thermodynamic identity for Φ is

$$d\Phi = -S dT - P dV - N d\mu. \quad (7.6)$$

The natural variables are the temperature, volume, and chemical potential. There are also three partial-derivative formula associated with it:

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

Using similar techniques in Section 5.2 and Section 6.5, there are two results respectively:

1. When T and μ are constant (thermal and diffusive equilibrium with a large reservoir), the grand free energy of a system tends to decrease:

$$dS_{\text{total}} = -\frac{1}{T} d\Phi.$$

2. The grand free energy is related to the grand partition function by

$$d\Phi = -kT \ln \mathcal{Z}. \quad (7.7)$$

Interestingly, when T and P are fixed, we already know that $G = \mu N$. Then the grand partition function is simply $\Phi = -PV$:

$$\Phi = U - TS - \mu N = U - TS + PV - PV - \mu N = G - PV - \mu N = -PV.$$

7.2 Bosons and Fermions

Gibbs factors are important to **quantum statistics**, where identical particles have a reasonable chance of wanting to occupy the same single-particle state. Recall that in Section 6.6, we assume that the system is not dense so that a system of N indistinguishable, noninteracting particles has the partition function

$$Z = \frac{1}{N!} Z_1^N, \quad (7.8)$$

where Z_1 is the partition function of a single particle. When the system is dense, this breaks down. We have introduced the problem with $1/N!$ in that section. Here is a more specific example: consider a system with two noninteracting particles with 5 states available: $\square\square\square\square\square$. (In this section the word “state” refers to a one-particle state, the state of the system is “the system state”.) Each can occupy one of the five states, all with energy 0 so every boltzmann factor is 1 and $Z = \Omega$. If the two particles are distinguishable, then the partition function is just $Z = 5 \times 5 = 25$. If they are *indistinguishable*, equation (7.8) predicts $Z = 5^2/2 = 12.5$. But Z is the same as the multiplicity Ω now, so it cannot be a half-integer.

Then how many system states are there? By brute force counting, we can easily get all system states, each specified by five numbers:

11000	01010	20000
10100	01001	02000
10010	00110	00200
10001	00101	00020
01100	00011	00002

similar to Einstein solids. There are 10 system states where two particles are at different states, and 5 system states where they are at the same state. If the two particles are distinguishable (**Maxwell-Boltzmann distribution**), then the first 10 system states are doubled, hence the total system states $10 \times 2 + 5$ as desired. If the two particles are indistinguishable, $1/N!$ correctly reduce 20 to 10, but it also incorrectly reduce 5 to 2.5. The last 5 system states are there no matter particles are distinguishable or indistinguishable.

However, the above example consider only identical particles that are *allowed* to be in the same state. These particles are called **bosons** (mediators like photons, pions, or nuclei like ^4He , etc.) There could be infinite number of bosons in a given state. Many

species of identical particles cannot occupy the same state, and they are called **fermions** (all elementary matter particles like electrons, quarks, neutrinos, or baryons like protons and neutrons, etc). This is known as the **Pauli exclusion principle**. If the identical particles are fermions, then the possible system states in the above analysis is only the first 10. Here is another way to tell bosons and fermions apart: just look at their spins (in units of \hbar). Particles with half-integer spins are fermions, while ones with integer spins are bosons. For example, an electron has spin $1/2$, so it is a fermion. A photon has spin 1 , so it is a boson. This rule can be derived from relativistic quantum mechanics, but is usually taken as an axiom in non-relativistic quantum mechanics.

As discussed in Boltzmann statistics, when the number of available single-particle states is much greater than the number of particles, $Z_1 \gg N$, the change of two particles occupying the same state is negligible. For an ideal gas, the quantum volume v_Q sets the scale, when

$$v_Q = \ell_Q^3 = \left(\frac{h^2}{2\pi m k T} \right)^{3/2} \ll V/N,$$

we are in the classical limit. There are several systems that violates $Z_1 \gg N$: systems that are extremely dense like neutron stars, or as cold as liquid helium, or composed of very light particles. This chapter deals with these extreme systems.

7.2.1 The Distribution Functions

For quantum statistics, the idea is to consider a system of one single-particle state instead of a particle. In other words, the system is a particular spatial wavefunction with a particular spin orientation. The reservoir consists of all the other possible single-particle states, or wavefunctions. Let the energy of a single-particle state occupied by one particle to be ϵ . Then when the state is unoccupied, the energy is 0, and when the state is occupied by n particles, the energy is $n\epsilon$. The probability of having n particles in this state is

$$\mathcal{P}(n) = \frac{1}{\mathcal{Z}} e^{-(n\epsilon - \mu n)/kT} = \frac{1}{\mathcal{Z}} e^{-n(\epsilon - \mu)/kT}.$$

If the particles are fermions, then n can only be 0 or 1 because no more than one fermion can occupy the same state. The grand partition function for fermions is thus

$$\mathcal{Z} = 1 + e^{-(\epsilon - \mu)/kT}.$$

The average number of particles in that state is called the **occupancy**:

$$\bar{n} = \sum_n n \mathcal{P}(n) = 0 \cdot \mathcal{P}(0) + 1 \cdot \mathcal{P}(1) = \frac{e^{-(\epsilon - \mu)/kT}}{1 + e^{-(\epsilon - \mu)/kT}} = \frac{1}{e^{(\epsilon - \mu)/kT} + 1}.$$

For fermions, the occupancy of a state is just the probability of the state being occupied. This formula is called the **Fermi-Dirac distribution**:

$$\boxed{\bar{n}_{\text{FD}} = \frac{1}{e^{(\epsilon - \mu)/kT} + 1}}. \quad (7.9)$$

It is a function of ϵ , μ , and T . Note that \bar{n}_{FD} goes to zero when $\epsilon \gg \mu$ and goes to 1 when $\epsilon \ll \mu$. Hence a state with energy much smaller than μ tends to be occupied. Moreover, it is easy to prove (and see on the graph) that \bar{n}_{FD} is “symmetric” about μ , with $\bar{n}_{\text{FD}}(\epsilon = \mu) = 1/2$.

If the particles are bosons, n can be any non-negative integer. The grand partition function is

$$\mathcal{Z} = 1 + e^{-(\epsilon - \mu)/kT} + e^{-2(\epsilon - \mu)/kT} + \dots = \sum_{n=0}^{\infty} \left[e^{-(\epsilon - \mu)/kT} \right]^n = \frac{1}{1 - e^{-(\epsilon - \mu)/kT}}.$$

The Gibbs factor cannot grows to infinity, so we require that $\mu < \epsilon$ and the series will converge. The average number of particles in a state is

$$\bar{n} = \sum_n n \mathcal{P}(n) = \sum_n n \frac{e^{-nx}}{\mathcal{Z}} = -\frac{1}{\mathcal{Z}} \sum_n \frac{\partial}{\partial x} e^{-nx} = -\frac{1}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial x},$$

where $x \equiv (\epsilon - \mu)/kT$. (This formula also works for fermions.) So the bosons have occupancy

$$\bar{n} = -(1 - e^{-x}) \frac{\partial}{\partial x} \frac{1}{1 - e^{-x}} = \frac{1}{e^{(\epsilon - \mu)/kT} - 1}.$$

This is called the **Bose-Einstein distribution**:

$$\boxed{\bar{n}_{\text{BE}} = \frac{1}{e^{(\epsilon - \mu)/kT} - 1}}. \quad (7.10)$$

It is also possible to derive the occupancy of the Boltzmann distribution. The probability of one particle being in a state of energy ϵ is

$$\mathcal{P}(s) = \frac{1}{Z_1} e^{-\epsilon/kT} \implies \bar{n}_{\text{Boltzmann}} = N\mathcal{P}(s) = \frac{N}{Z_1} e^{-\epsilon/kT}.$$

We can relate N/Z_1 with chemical potential μ by (6.8), $\mu = -kT \ln(Z_1/N)$, so

$$\bar{n}_{\text{Boltzmann}} = e^{\mu/kT} e^{-\epsilon/kT} = e^{-(\epsilon-\mu)/kT}.$$

We can summarize these occupancies (or **distribution functions**) in one equation:

$$\bar{n} = \frac{1}{e^{(\mu-\epsilon)/kT} + \alpha} \quad (7.11)$$

where $\alpha = \pm 1, 0$. We have 1 for fermion, -1 for boson, and 0 for classical (Boltzmann). Figure 7.1 is a plot of the three $\bar{n}(\epsilon)$ at fixed temperature and chemical potential.

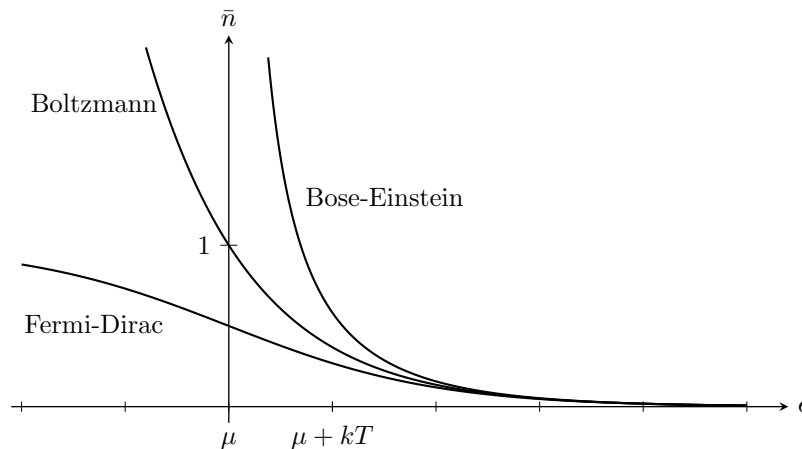


Figure 7.1: Comparison of the Fermi-Dirac, Bose-Einstein, and Boltzmann distributions, all for the same value of μ . When $(\epsilon - \mu)/kT \gg 1$, the three distributions become equal.

7.3 Degenerate Fermi Gas

A gas of fermions at very low temperature is an application of quantum statistics and the Fermi-Dirac distribution. One typical example is the conduction electrons inside a metal. When the temperature is sufficiently low such that $V/N \ll v_Q$, we need to use quantum statistics. For an electron at room temperature, the quantum volume is about

$$v_Q = \left(\frac{h^2}{2\pi m kT} \right)^{3/2} \approx (4.3 \text{ nm})^3.$$

However, inside a metal, the volume per electron V/N is only about $(0.2 \text{ nm})^3$, which is much smaller than the quantum volume. The room temperature is too low to apply Boltzmann statistics to conduction electrons. For now, we shall pretend the temperature is at $T = 0$, and later we will generalize low temperature cases.

7.3.1 Zero Temperature

At $T = 0$, the Fermi-Dirac distribution becomes a step function:

$$\bar{n} = \frac{1}{e^{(\epsilon-\mu)/kT} + 1} \longrightarrow \bar{n} = \begin{cases} 1, & \epsilon < \mu \\ 0, & \epsilon > \mu. \end{cases}$$

In this case we define μ as the **Fermi energy**, ϵ_F :

$$\epsilon_F \equiv \mu(T = 0).$$

When a Fermi gas is so cold that nearly all states with $\epsilon < \epsilon_F$ are occupied and nearly all states with $\epsilon > \epsilon_F$ are unoccupied, it is said to be **degenerate**. (This name is totally unrelated to the one in quantum mechanics.) The value of ϵ_F is determined by the number of electrons in the gas. Suppose you are adding electrons to a box one at a time. The electrons will fill up the energy levels due to Pauli exclusion principle, and the last electron will have energy just below ϵ_F . To add one more electron, you need to put in the energy $\epsilon_F = \mu$. This is perfectly consistent with the definition $\mu = (\partial U / \partial N)_{S,V}$.

Now consider free electrons (electrostatic force ignored) inside a box of $V = L^3$. According to quantum mechanics in 3D, electrons can have momentum

$$p_x = \frac{hn_x}{2L}, \quad p_y = \frac{hn_y}{2L}, \quad p_z = \frac{hn_z}{2L}, \quad \text{where } n_x, n_y, n_z = 1, 2, 3, \dots$$

and energy levels

$$\epsilon = \frac{p^2}{2m} = \frac{h^2}{8mL^2}(n_x^2 + n_y^2 + n_z^2).$$

We can specify the allowed energies in the “ n -space”, where the three axes are labeled n_x , n_y , and n_z , and each state is represented by the vector $\mathbf{n} = (n_x, n_y, n_z)$. Note that in this space, the energy is proportional to the square of the distance from the origin, $n_x^2 + n_y^2 + n_z^2$. When we add electrons into the box, we are filling up the n -space starting from the center, until all energy levels are filled. All states live within the first octant of a sphere with radius n_{\max} , which specifies highest energy states. The Fermi energy is thus

$$\epsilon_F = \frac{h^2 n_{\max}^2}{8mL^2}.$$

The total number of states within the box is the volume of 1/8 (first octant) of the sphere, times 2 (this accounts for two spin orientations of electrons), so

$$N = 2 \times (\text{volume of eighth-sphere}) = 2 \cdot \frac{1}{8} \cdot \frac{4}{3} \pi n_{\max}^3 = \frac{\pi n_{\max}^3}{3}.$$

These two equations give the Fermi energy of the gas as a function of total electrons N and the volume $V = L^3$,

$$\epsilon_F = \frac{h^2}{8m} \left(\frac{3N}{\pi V} \right)^{2/3}. \quad (7.12)$$

Note that the Fermi energy depends only on the number density of electrons N/V , so it is an intensive quantity. The total energy of the system is just the triple sum, which can be approximated as an integral,

$$\begin{aligned} U &= 2 \sum_{n_x, n_y, n_z} \epsilon(\mathbf{n}) \approx 2 \int \epsilon(\mathbf{n}) dn_x dn_y dn_z \\ &= 2 \int_0^{n_{\max}} dn \int_0^{\pi/2} d\theta \int_0^{\pi/2} d\phi n^2 \epsilon(n) \sin \theta \\ &= \pi \int_0^{n_{\max}} \epsilon(n) n^2 dn = \frac{\pi h^2}{8mL^2} \int_0^{n_{\max}} n^4 dn = \frac{\pi h^2 n_{\max}^5}{40mL^2}. \end{aligned}$$

We use spherical coordinates in n -space and multiplying by 2 accounts for the spin orientations of electrons. The final answer can be written in terms of the Fermi energy and total number of electrons:

$$U = \frac{3}{5} N \epsilon_F \quad (7.13)$$

This means the average energy of electrons is 3/5 the Fermi energy. Though the above analysis is based on electrons in a box, the result works for any Fermi gas in any volume at $T = 0$. Plugging in some numbers, the Fermi energy is $\epsilon_F \sim \text{eV}$, while the average thermal energy of a particle at room temperature is only $kT \approx 1/40 \text{ eV}$. Thus there are two equivalent statements for degenerate Fermi gases:

$$\frac{V}{N} \ll v_Q \iff kT \ll \epsilon_F.$$

The approximation $T \approx 0$ is quite accurate in many situations. The temperature that a Fermi gas would have such that $kT = \epsilon_F$ is called the **Fermi temperature**: $T_F \equiv \epsilon_F/k$. We can calculate the pressure of a degenerate Fermi gas using $P = -(\partial U / \partial V)_{S,V}$:

$$P = -\frac{\partial}{\partial V} \left[\frac{3}{5} \frac{Nh^2}{8m} \left(\frac{3N}{\pi V} \right)^{2/3} \right] = \frac{2N\epsilon_F}{5V} = \frac{2}{3} \frac{U}{V} \quad (7.14)$$

This is called the **degeneracy pressure** of a Fermi gas. Such pressure has nothing to do with electrostatic or other kinds of fundamental forces. It comes entirely from Pauli exclusion principle, and it is what keeps electrons from falling into protons by electrostatic force.

7.3.2 Small Nonzero Temperatures

The Fermi gas with zero temperature assumption cannot provide any information about the heat capacity. We shall see how nonzero but small temperatures would affect a Fermi gas.

At temperature T , any particle may acquire a thermal energy of roughly kT . However, most of the electrons cannot acquire such amount of energy because the higher energy states are already occupied. Only those with energy about $\epsilon_F - kT$ can jump to the next level. (This also leave some space for lower energy electrons.) Because of the energy tolerance kT , the number of electrons that are affected by an increase in T is proportional to T . Meanwhile, this number should be proportional to N as well because it is an extensive quantity. These arguments give the increase in energy of a Fermi gas when temperature goes from zero to T ,

$$\begin{aligned}\Delta E &\propto (\text{number of affected electrons}) \times (\text{energy acquired by each}) \\ &\propto (NkT) \times (kT) \\ &\propto N(kT)^2.\end{aligned}$$

By dimensional analysis, the proportionality constant should have units of $(\text{energy})^{-1}$. The only constant available in this context is ϵ_F , so the increase in energy must be some dimensionless constant times $N(kT)^2/\epsilon_F$. This dimensionless constant is determined to be $\pi^2/4$. In conclusion, the total energy of a degenerate Fermi gas for $kT \ll \epsilon_F$ and the heat capacity are

$$U = \frac{3}{5}N\epsilon_F + \frac{\pi^2}{4} \frac{N(kT)^2}{\epsilon_F}, \quad C_V = \left(\frac{\partial U}{\partial T} \right)_V = \frac{\pi^2 N k^2 T}{2\epsilon_F}. \quad (7.15)$$

7.3.3 The Density of States

To quantitatively determine the energy of a degenerate Fermi gas, consider the change of variable from n to the electron energy ϵ :

$$\epsilon = \frac{h^2}{8mL^2}n^2, \quad n = \sqrt{\frac{8mL^2}{h^2}}\sqrt{\epsilon}, \quad dn = \sqrt{\frac{8mL^2}{h^2}} \frac{1}{2\sqrt{\epsilon}} d\epsilon.$$

The energy integral at $T = 0$ becomes

$$U = \int_0^{\epsilon_F} \epsilon \left[\frac{\pi}{2} \left(\frac{8mL^2}{h^2} \right)^{3/2} \sqrt{\epsilon} \right] d\epsilon.$$

The term in the square bracket is called the **density of states**, denotes as $g(\epsilon)$,

$$g(\epsilon) = \frac{\pi(8m)^{3/2}}{2h^3} V \sqrt{\epsilon} = \frac{3N}{2\epsilon_F^{3/2}} \sqrt{\epsilon}. \quad (7.16)$$

It indicates the number of states within the energy interval $[\epsilon, \epsilon + d\epsilon]$. Note that the density of states does not depend on N , where in the second equality N is canceled by $\epsilon_F^{3/2}$. It is a useful quantity to be integrated. If you want the number of states between ϵ_1 and ϵ_2 , just integrate $g(\epsilon)$ in this range. Moreover, the density of states is a pure quantum mechanical quantity, and of course different models would have different forms of $g(\epsilon)$. Once the density of states of some system is known, it can be applied to thermal physics.

For an electron gas at zero temperature, the total number of electrons is just

$$N = \int_0^{\epsilon_F} g(\epsilon) d\epsilon.$$

But when $T \neq 0$, not all states are occupied according to the previous section. In fact, we already know the probability of a state with energy ϵ being occupied—the Fermi-Dirac distribution. Therefore, we will modify the integrals of total number of electrons and total energy to

$$\begin{aligned}N &= \int_0^\infty g(\epsilon) \bar{n}_{\text{FD}}(\epsilon) d\epsilon = \int_0^\infty \frac{g(\epsilon)}{e^{(\epsilon-\mu)/kT} + 1} d\epsilon. \\ U &= \int_0^\infty \epsilon g(\epsilon) \bar{n}_{\text{FD}}(\epsilon) d\epsilon = \int_0^\infty \frac{\epsilon g(\epsilon)}{e^{(\epsilon-\mu)/kT} + 1} d\epsilon.\end{aligned}$$

Note that the chemical potential is no longer equal to the Fermi energy when $T \neq 0$, but is slightly smaller instead. Figure 7.2 shows the graph of the integrand $g(\epsilon)\bar{n}_{\text{FD}}$ of the N -integral.

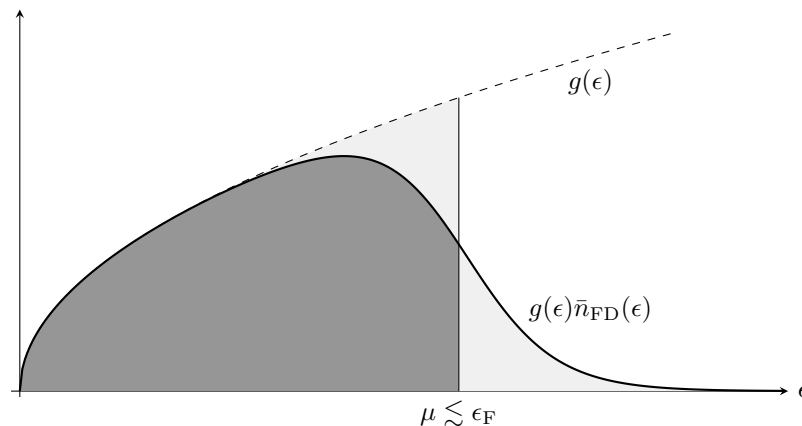


Figure 7.2: At nonzero T , the number of fermions per unit energy is given by the density of states times the Fermi-Dirac distribution. Because increasing the temperature does not change the total number of fermions, the two lightly shaded areas must be equal. Since $g(\epsilon)$ is greater above ϵ_F than below and the Fermi-Dirac distribution is symmetric about μ , this means that the chemical potential decreases as T increases. This graph is drawn for $T/T_F = 0.1$; at this temperature μ is about 1% less than ϵ_F .

The chemical potential now plays a role of fixing the total number of particles. Once we get N in a closed form from the integral, it is possible to solve for $\mu(T)$. After obtaining $\mu(T)$, plug it into the U -integral gives the energy as a function of T , and hence the heat capacity. Unfortunately, these integrals do not have a closed form. They can only be evaluated by approximation in the limit $kT \ll \epsilon_F$.

7.3.4 The Sommerfeld Expansion

The **Sommerfeld expansion** can tell us the answer of $\mu(T)$ and $U(T)$ in the limit $kT \ll \epsilon_F$:

$$\frac{\mu}{\epsilon_F} = 1 - \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 + \dots, \quad (7.17)$$

$$U = \frac{3}{5} N \epsilon_F + \frac{\pi^2}{4} \frac{N (kT)^2}{\epsilon_F} + \dots \quad (7.18)$$

For proofs, see Appendix A.3.1. There emerges the $\pi^2/4$ coefficient for the energy correction. Also note that the chemical potential decreases as T increases.

7.4 Blackbody Radiation

7.4.1 The Planck Distribution

Classical electromagnetism predicts electromagnetic waves as propagating fields through all space. If those fields are confined to a box with fixed temperature (like in an oven), they are actually infinite superpositions of standing waves with different frequencies. Each standing wave is like an harmonic oscillator: there are two quadratic degrees of freedom, so each has an energy of kT . Since there are infinite number of standing waves, the energy within the box is infinite, which obviously disagrees with experiments. This disagreement is called the **ultraviolet catastrophe**. Historically, it led to the birth of quantum mechanics.

In quantum mechanics, a harmonic oscillator cannot have arbitrary frequency or energy. The allowed energies are $E_n = (n + 1/2)\hbar\omega$. The partition function for one oscillator is

$$Z = 1 + e^{-\beta\hbar\omega} + e^{-2\beta\hbar\omega} + \dots = \frac{1}{1 - e^{-\beta\hbar\omega}}.$$

This gives the average energy,

$$\langle E \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}.$$

If we see the energy as in units of $h\nu = \hbar\omega$ (conventional notation), then the average number of units of energy in an oscillator is

$$\bar{n}_{\text{Pl}} = \frac{1}{e^{h\nu/kT} - 1}. \quad (7.19)$$

This is called the **Planck distribution**. When $h\nu \gg kT$, the probability is exponentially suppressed. By making use of the quantized energy, the ultraviolet catastrophe is resolved.

The “units” of energy can be thought of as particles called **photons**. Photons are bosons, so they should obey the Bose-Einstein distribution:

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

With $\epsilon = h\nu$, the Planck distribution (7.19) says that the chemical potential for photons should be $\mu = 0$. There are two explanations for this:

1. At thermal equilibrium with T and V fixed, we know that the Helmholtz free energy is minimized. Since photons number is not conserved (as it can be absorbed by an electron), a system of N photons would take a specific N such that F is minimized. Then at equilibrium, by the partial derivative relation of F and μ :

$$\left(\frac{\partial F}{\partial N} \right)_{T,V} = \mu = 0.$$

2. Consider the reaction with a photon (γ) is emitted or absorbed by an electron:

$$e^- \longleftrightarrow e^- + \gamma.$$

The equilibrium condition is $\mu_e = \mu_e + \mu_\gamma$, so the chemical potential for photons is zero.

7.4.2 Total Particle Number and Energy

The Planck distribution provides the occupancy of photons. We shall compute the total number of photons in a box and the total energy of the system. Again, consider first the “particle in a 1D box” of length L . Photons have momentum

$$p = \frac{hn}{2L}, \quad n = 1, 2, 3, \dots$$

Unlike electrons or other particles, photons are relativistic—their energies are given by

$$\epsilon = pc = \frac{hcn}{2L}.$$

Generalizing to a 3D box, the energy is

$$\epsilon = c\sqrt{p_x^2 + p_y^2 + p_z^2} = \frac{hc}{2L}\sqrt{n_x^2 + n_y^2 + n_z^2} = \frac{hcn}{2L}, \quad n_x, n_y, n_z = 1, 2, 3, \dots$$

The next step is similar to the one in the last section. Summing over n , the total number of particles and energy are

$$N = 2 \sum_{n_x, n_y, n_z} \bar{n}_{\text{Pl}}(\epsilon) \approx 2 \int_0^\infty dn \int_0^{\pi/2} d\theta \int_0^{\pi/2} d\phi \frac{n^2 \sin \theta}{e^{hcn/2LkT} - 1}.$$

$$U = 2 \sum_{n_x, n_y, n_z} \epsilon(n) \bar{n}_{\text{Pl}}(\epsilon) \approx \int_0^\infty dn \int_0^{\pi/2} d\theta \int_0^{\pi/2} d\phi \frac{hcn}{L} \frac{n^2 \sin \theta}{e^{hcn/2LkT} - 1}.$$

The factor of 2 accounts for the two polarizations of photons. Note that this time n runs from zero to infinity. Using the substitution $x = hcn/2LkT$, the N -integral is

$$N = \pi \int_0^\infty \frac{n^2}{e^{hcn/2LkT} - 1} dn = \pi \left(\frac{2LkT}{hc} \right)^3 \int_0^\infty \frac{x^2}{e^x - 1} dx = 8\pi V \left(\frac{kT}{hc} \right)^3 \int_0^\infty \frac{x^2}{e^x - 1} dx.$$

This integral does not have an analytical solution, the numerical value is $2\zeta(3) \approx 2.404$ where $\zeta(z)$ is the Riemann-zeta function.

The Planck Spectrum

Changing variables to $\epsilon = hc\nu/2L = h\nu$, the U -integral becomes

$$\frac{U}{V} = \int_0^\infty \frac{8\pi\epsilon^3/(hc)^3}{e^{\epsilon/kT} - 1} d\epsilon = \int_0^\infty \frac{8\pi h\nu^3/c^3}{e^{h\nu/kT} - 1} d\nu.$$

The integrand is interpreted as the energy density per photon energy, or the **spectrum** of the photons:

$$u(\epsilon) = \frac{8\pi}{(hc)^3} \frac{\epsilon^3}{e^{\epsilon/kT} - 1}, \quad u(\nu) = \frac{8\pi h}{c^3} \frac{\nu^3}{e^{h\nu/kT} - 1}. \quad (7.20)$$

To evaluate the integral analytically, make a substitution $x = \epsilon/kT$,

$$\frac{U}{V} = \frac{8\pi(kT)^4}{(hc)^3} \int_0^\infty \frac{x^3}{e^x - 1} dx.$$

Note that the U -integral can also be written as

$$U = \int_0^\infty \epsilon \left[\frac{8\pi V \epsilon^2}{(hc)^3} \right] \bar{n}_{\text{BE}}(\epsilon) d\epsilon.$$

The term in the square bracket is thus the density of states for photon (or other kinds of ultra-relativistic particles):

$$g(\epsilon) = \frac{8\pi V \epsilon^2}{(hc)^3}. \quad (7.21)$$

Back to the spectrum (7.20). By numerical calculation, the spectrum peaks at $\epsilon \approx 2.82kT$. **Wien's law** says that higher temperature gives a higher photon energy. Figure 7.3 shows the integrand as a function of $x = \epsilon/kT$.

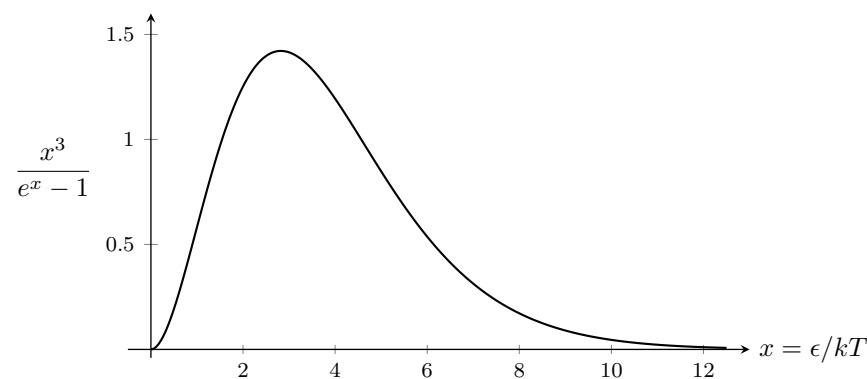


Figure 7.3: The Planck spectrum, plotted in terms of the dimensionless variable $x = \epsilon/kT = h\nu/kT$. The area under any portion of this graph, multiplied by $8\pi(kT)^4/(hc)^3$, equals the energy density of electromagnetic radiation within the corresponding frequency (or photon energy) range.

The total energy is just the integral. It has an analytical solution:

$$\frac{U}{V} = \frac{8\pi(kT)^4}{(hc)^3} \int_0^\infty \frac{x^3}{e^x - 1} dx = \frac{8\pi^5(kT)^4}{15(hc)^3}. \quad (7.22)$$

The total energy is proportional to the *fourth* power of the temperature. It is also possible to the spectrum in terms of wavelength $\lambda = hc/\epsilon$ (see Appendix A.3.3):

$$u(\lambda) = \frac{8\pi hc}{\lambda^5} \frac{1}{e^{hc/kT\lambda} - 1}. \quad (7.23)$$

It peaks at $\lambda = hc/4.97kT$. One may ask: why is the peak not at $\lambda_{\text{peak}} = hc/\epsilon_{\text{peak}} = hc/2.82kT$? Let's look at the two photon spectrum: (7.20) shows the energy density per unit photon energy ϵ , while (7.23) shows the energy density per unit wavelength λ . The problem is the photon energy and the wavelength are related to each other in a nonlinear way. For example, consider a 1 eV of photon energy unit $\delta\epsilon$. If we choose the photon energy range from 1 eV to 2 eV, this corresponds to a wavelength range of 621 nm to 1242 nm. If it is from 11 eV to 12 eV, the corresponding wavelength range is from 104 nm to 112 nm, which is a much smaller $\delta\lambda$.

Other Quantities

For a box of photons with volume V , the heat capacity is

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V = 4aVT^3,$$

where $a = 8\pi^5 k^4 / 15(hc)^3$. This expression works for any temperature, so we can calculate the entropy with it:

$$S(T) = \int_0^T \frac{C_V(T')}{T'} dT' = 4aV \int_0^T (T')^2 dT' = \frac{4}{3}aVT^3 = \frac{32\pi^5}{45} V \left(\frac{kT}{hc} \right)^3 k. \quad (7.24)$$

Note that the entropy of photons is the same as the total number of photons. They only differ by a dimensionless constant and a Boltzmann constant.

The pressure of a photon gas is $1/3$ times the energy density U/V . This can be obtained by using $P = -\left(\frac{\partial U}{\partial V}\right)_{S,N}$. To do this, we need to control S (which is the same as controlling N) and eliminate the uncontrollable T . First let $a = 8\pi^5 k^4 / 15(hc)^3$, and

$$U = aVT^4 \quad \text{and} \quad S = \frac{4}{3}aVT^3.$$

Solving for T and substitute into U ,

$$U = aV \left(\frac{3S}{4aV} \right)^{4/3} = \left(\frac{3S}{4} \right)^{4/3} (aV)^{-1/3}.$$

The pressure is

$$P = -\left(\frac{\partial U}{\partial V} \right)_S = \frac{1}{3} \left(\frac{3S}{4} \right)^{4/3} a^{-1/3} V^{-4/3} = \frac{1}{3} \frac{U}{V} = \frac{1}{3} aT^4. \quad (7.25)$$

The constant a is often called the **radiation constant**.

Once we have the internal energy and entropy, we can calculate the Helmholtz free energy,

$$F = U - TS = aVT^4 - T \cdot \frac{4}{3}aVT^3 = -\frac{1}{3}aVT^4 = -\frac{1}{3}U.$$

You may check that $S = -(\partial F / \partial T)_V$ and $P = -(\partial F / \partial V)_T$ agrees with the above results. It is also possible to calculate F directly from the partition function by summing over modes (see Appendix A.3.2)

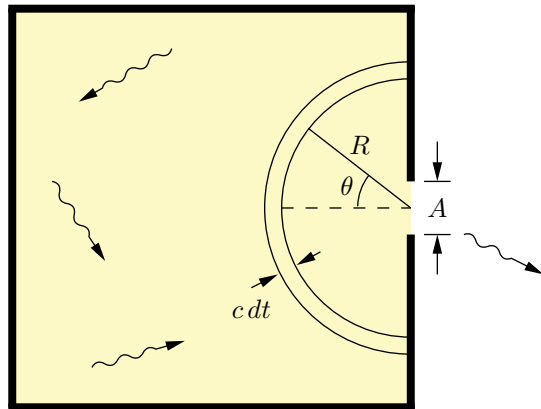
7.4.3 The Cosmic Background Radiation

The grandest photon gas is the cosmic background radiation. This radiation fills the entire observable universe with temperature (nowadays) 2.73 K. It is believed that these photons were created in the early universe when the temperature is ~ 3000 K. Since the universe is expanding, the wavelengths of photons are stretched out and thus have less energy. The spectrum peaks at $\epsilon = 2.82kT \approx 6.6 \times 10^{-4}$ eV in the far infrared. They are hard to detect because they do not penetrate the atmosphere. Fortunately, the microwave region of the spectrum are detectable, so we name this radiation the **cosmic microwave background**.

7.4.4 Stefan-Boltzmann Law

In previous sections we only consider photons inside a box in thermal equilibrium. We haven't truly talk about blackbody radiation, or how photons are emitted by a hot object.

Consider a box of photon gas, but with a small hole of area A . Photons have the same speed in vacuum, so low-energy photons will have the same probability escaping the hole as energetic photons. The spectrum inside the box should be the same as the one outside. We want to know the total amount of radiation coming out. Suppose there are some photons that can escape through the hole within a time interval dt . We know that for a time t prior to their escape, they were once located at some positions within a shell of radius $R = ct$ and thickness $c dt$.



The volume element of the shell is

$$dV = R^2 \sin \theta d\theta d\phi (c dt).$$

The energy in this volume element is

$$dE = \frac{U}{V} c R^2 \sin \theta d\theta d\phi dt.$$

However, not all energy in this volume will escape the hole because most of the photons are pointing in the wrong direction. The probability of pointing in the right direction is the apparent area of the hole seen from the volume, divided by the total possible area of the flow:

$$\mathcal{P} = \frac{A \cos \theta}{4\pi R^2}.$$

The true energy escaping from the volume element is

$$dE_{\text{esc}} = \frac{A \cos \theta}{4\pi} \frac{U}{V} c \sin \theta d\theta d\phi dt.$$

The power escaping is then

$$\frac{dE}{dt} = \int_0^{2\pi} d\phi \int_0^{\pi/2} d\theta \frac{A \cos \theta}{4\pi} \frac{U}{V} c \sin \theta = \frac{A U}{4} \frac{c}{V}.$$

From this we can obtain the flux, or power per unit area:

$$F_{\text{rad}} = \frac{c U}{4 V}.$$

Plug in (7.22), we obtain the **Stefan-Boltzmann law**:

$$F_{\text{rad}} = \frac{2\pi^5 (kT)^4}{15 h^3 c^2} = \sigma T^4, \quad (7.26)$$

where $\sigma = 2\pi^5 k^4 / 15 h^3 c^2 \approx 5.67 \times 10^{-8} \text{ W/m}^2 \text{ K}^4$ is known as the **Stefan-Boltzmann constant**. The Stefan-Boltzmann law not only describes photons leaking through a hole, but also applies to radiation from non-reflecting surfaces at temperature T . This kind of radiation is called the **blackbody radiation**.

Here is the reasoning. Suppose there is a perfect blackbody surface that matches the size of the hole of the box. The photon gas in the box are at the same temperature as the blackbody. Then they must emit the same power. If one of them radiates more, then the other will heat up *when they are at thermal equilibrium*! That violates the second law of thermodynamics. Therefore, the blackbody must emit the same power as the photon gas does. Furthermore, if we have some filter that only allow a certain wavelength to pass through, then the same reasoning above holds for that wavelength. This is saying that the *whole* spectrum of the blackbody is the same as that of the hole.

For an object that is not a perfect blackbody, we define its **emissivity** e to be the fraction of photons absorbed. To remain in thermal equilibrium with the hole, the blackbody should also emit only e fraction of photons. The emissivity equals 1 for a perfect blackbody, and 0 for a perfect reflector. The power of an object with surface area A and emissivity e is $\sigma e A T^4$.

7.5 Debye Theory of Solids

In Chapter 2 we spent lots of time on the Einstein model of a solid. Each atom acts as an independent three-dimensional harmonic oscillator. Appendix A.2.1 derives the heat capacity of N harmonic oscillators, so for N atoms (that is $3N$ oscillators), the heat capacity is

$$C_V = 3Nk \frac{(\epsilon/kT)^2 e^{\epsilon/kT}}{(e^{\epsilon/kT} - 1)^2}.$$

When $kT \gg \epsilon$, the heat capacity is $3Nk$, consistent with the equipartition theorem. It also predicts that the heat capacity falls off exponentially to zero as $T \rightarrow 0$. However, experiments show that $C_V \propto T^3$ at low temperature for real solids.

You may have already noticed the problem with Einstein solids. This model is over-simplified because atoms do not vibrate independently. There are both low-frequency modes and high-frequency modes with different energy units. At low temperatures where high-frequency modes freeze out, but some low-frequency modes are still vibrating. Therefore, the heat capacity does not fall off that dramatically as the Einstein model predicts.

Oscillation modes of solids are similar to electromagnetic modes in the last section. We may want to solve the problem using a similar approach. In the following analysis we will just call mechanical oscillations “sound waves”. Here are a few differences to notice between sound waves and EM waves:

- Sound waves are much slower and they depend on the property of materials. Its speed is denoted as c_s , and we will take it as a constant through the material.
- Sound waves can be polarized both transversely and longitudinally. The longitudinally polarized waves add one more polarization. We assume that all three polarizations have the same speed.
- Sound waves cannot have arbitrarily short wavelengths. Their wavelength must be at least twice the atomic spacing.

The energy of one mode of oscillation is also equally spaced,

$$\epsilon = h\nu = \frac{hc_s}{\lambda} = \frac{hc_s n}{2L},$$

where L is the length of the solid and $n = |\mathbf{n}|$ is the magnitude of the n -space vector. At temperature T , the number of units of energy a mode contains is given by the Planck distribution,

$$\bar{n}_{\text{Pl}} = \frac{1}{e^{\epsilon/kT} - 1}.$$

We call these “particles” **phonons**. Again, they have zero chemical potential like photons. The total energy of the solid is given by

$$U = 3 \sum_{n_x} \sum_{n_y} \sum_{n_z} \epsilon \bar{n}_{\text{Pl}}(\epsilon).$$

Unlike photons, the n cannot be arbitrarily large. Consider atoms lining up in a one-dimensional row. The n th oscillation mode should have n crests and troughs in total. It would be nonsense if n exceeds the number of atoms in that row. If the solid is a perfect cube, then the number of atoms in each direction is $\sqrt[3]{N}$. This should be the upper limit of each sum.

In general, the solid is not a perfect cube. But note that the energy depends only on $|\mathbf{n}|$, so we do the usual trick: do not care about the volume and integrate over the first octant of a sphere in n -space. To preserve the total number of degrees of freedom, Peter Debye chose the volume of the eighth-sphere to be N . This gives

$$N = \frac{1}{8} \frac{4\pi n_{\text{max}}^3}{3} \implies n_{\text{max}} = \left(\frac{6N}{\pi} \right)^{1/3}.$$

It turns out that Debye’s approximation is exact in both the high- and low-temperature limits, and a pretty good one in intermediate temperatures. The total energy turns to an integral,

$$U = 3 \int_0^{n_{\text{max}}} dn \int_0^{\pi/2} d\theta \int_0^{\pi/2} d\phi n^2 \sin \theta \frac{\epsilon}{e^{\epsilon/kT} - 1} = \frac{3\pi}{2} \int_0^{n_{\text{max}}} \frac{hc_s}{2L} \frac{n^3 dn}{e^{hc_s n/2LkT} - 1}.$$

Making the change of variable $x = hc_s n / 2LkT$, the upper limit of the integral becomes

$$x_{\max} = \frac{hc_s n_{\max}}{2LkT} = \frac{hc_s}{2kT} \left(\frac{6N}{\pi V} \right)^{1/3} \equiv \frac{T_D}{T},$$

where T_D is the **Debye temperature**. After the change of variable, the U -integral is

$$U = \frac{9NkT^4}{T_D^3} \int_0^{T_D/T} \frac{x^3}{e^x - 1} dx.$$

This integral cannot be done analytically, but it is valid for all temperatures.

We can check the high- and low-temperature limits. When $T \gg T_D$, the variable x is always small, so we can approximate $e^x = 1 + x$. The integral is just

$$U = \frac{9NkT^4}{T_D^3} \int_0^{T_D/T} x^2 dx = \frac{9NkT^4}{T_D^3} \left[\frac{1}{3} \left(\frac{T_D}{T} \right)^3 \right] = 3NkT.$$

It agrees with the equipartition theorem. The heat capacity is $C_V = 3Nk$. When $T \ll T_D$, the upper limit of the integral is a large number. Because of the e^x in the denominator, the integrand is almost zero for large x , so we can just set the upper limit to be infinity. Then the integral looks a lot familiar—it is $\pi^4/15$. The total energy at low-temperature limit is

$$U = \frac{3\pi^4}{5} \frac{NkT^4}{T_D^3}. \quad (7.27)$$

Differentiating with respect to T , the heat capacity is

$$C_V = \frac{12\pi^4}{5} \left(\frac{T}{T_D} \right)^3 Nk.$$

It agrees with the experimental result $C_V \propto T^3$ at low temperature. For metals, they have conduction electrons (as a Fermi gas). They contribute to the total heat capacity of a linear term according to (7.15), so the total heat capacity is

$$C = \gamma T + \frac{12\pi^4 Nk}{5T_D^3} T^3,$$

where $\gamma = \pi^2 Nk^2 / 2\epsilon_F$.

7.6 Bose-Einstein Condensation

Not all bosons are photons that can be created or destroyed easily. Atoms or nuclei with integer spins (such as ^4He with spin 0) are also bosons, and their chemical potential may not be zero. In particular, we *need* to determine the chemical potential to apply the Bose-Einstein distribution. At the end, we will find marvelous phenomena of a gas of bosons when the temperature is below a critical value.

At zero temperature, all bosonic atoms will be in the ground state. If they are in a box of $V = L^3$, they have energy

$$\epsilon_0 = \frac{h^2}{8mL} (1^2 + 1^2 + 1^2) = \frac{3h^2}{8mL^2},$$

which is a very small energy. At any temperature, the number of atoms in the ground state is given by the Bose-Einstein distribution:

$$N_0 = \frac{1}{e^{(\epsilon_0 - \mu)/kT} - 1}.$$

If the temperature is sufficiently low, N_0 will be large, $e^{(\epsilon_0 - \mu)/kT}$ will be very close to 1, and $(\epsilon_0 - \mu)/kT$ will be small. Thus, we can Taylor expand the exponential to get

$$N_0 = \frac{1}{1 + (\epsilon_0 - \mu)/kT - 1} = \frac{kT}{\epsilon_0 - \mu} \quad [N_0 \gg 1].$$

The chemical potential then must be ϵ_0 when $T = 0$, and slightly smaller than ϵ_0 when T is low. The question is: how low should the temperature be to have a large N_0 ? And how should we determine the chemical potential?

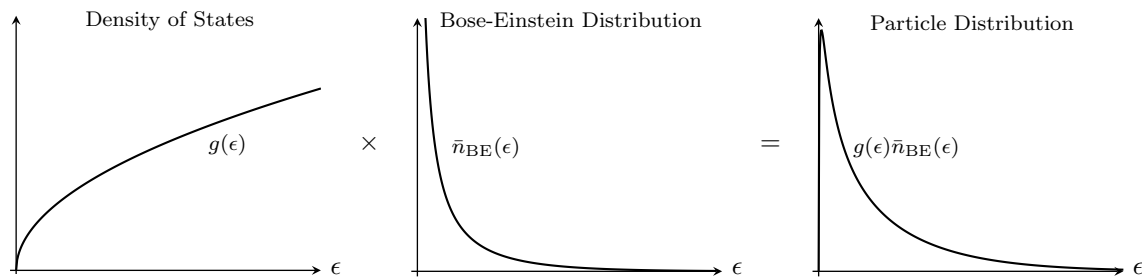


Figure 7.4: The distribution of bosons as a function of energy is the product of two functions, the density of states and the Bose-Einstein distribution. The chemical potential is drawn slightly less than zero.

In general, μ is like a normalization condition so that the sum of all states is the total number of atoms N ,

$$N = \sum_s \frac{1}{e^{(\epsilon_s - \mu)/kT} - 1}.$$

We can repeat this normalization process for all T and get the chemical potential. It is easier to make use of the density of states and write N as an integral,

$$N = \int_0^\infty g(\epsilon) \frac{1}{e^{(\epsilon - \mu)/kT} - 1} d\epsilon. \quad (7.28)$$

This approximation requires $kT \gg \epsilon_0$. The density of states for spin-zero bosons is (7.16) divided by 2 because there is only one spin orientation,

$$g(\epsilon) = \frac{2}{\sqrt{\pi}} \left(\frac{2\pi m}{h^2} \right)^{3/2} V \sqrt{\epsilon}. \quad (7.29)$$

The density of states, Bose-Einstein Distribution, and their product are plotted in Figure 7.4. The N -integral cannot be done analytically, so we need to guess a value of μ that works to match the total particle number. One may take $\mu = 0$ (because it is the easiest) for the first guess. By changing variables $x = \epsilon/kT$ and plugging in $\mu = 0$, the N -integral becomes

$$N = \frac{2}{\sqrt{\pi}} \left(\frac{2\pi m}{h^2} \right)^{3/2} V \int_0^\infty \frac{\sqrt{\epsilon} d\epsilon}{e^{\epsilon/kT} - 1} = \frac{2}{\sqrt{\pi}} \left(\frac{2\pi m kT}{h^2} \right)^{3/2} V \int_0^\infty \frac{\sqrt{x} dx}{e^x - 1}.$$

The integral is about 2.315. Combining it with $2/\sqrt{\pi}$ gives

$$N = 2.612 \left(\frac{2\pi m kT}{h^2} \right)^{3/2} V.$$

With V constant, we get the total particle number that depends on the temperature even for atoms, which is obviously wrong. There can be only one (and should be one) temperature that makes N correct. We will define it to be T_c ,

$$N = 2.612 \left(\frac{2\pi m kT_c}{h^2} \right)^{3/2} V, \quad kT_c = 0.527 \left(\frac{h^2}{2\pi m} \right) \left(\frac{N}{V} \right)^{2/3}.$$

7.6.1 The Condensation Temperature

When $T > T_c$, the chemical potential should be smaller to maintain the same N according to (7.28). When $T < T_c$, we need to think more about whether replacing the discrete sum by the integral is valid. The integrand of the N -integral is proportional to $\sqrt{\epsilon}/\epsilon$ as ϵ goes to zero, i.e. it blows up at $\epsilon = 0$ when $\mu = 0$. We are not sure whether this infinite spike correctly represents the sum. We already know that when $\mu \approx 0$ and T small enough, the number of atoms in the ground state should be huge. The integral does not include this. Instead, if we compute the integral with the lower limit $\epsilon^+ > \epsilon_0$ but much smaller than kT , we would also get

$$N = \frac{2}{\sqrt{\pi}} \left(\frac{2\pi m kT}{h^2} \right)^{3/2} V \int_{\epsilon^+/kT}^\infty \frac{\sqrt{x} dx}{e^x - 1} \approx 2.612 \left(\frac{2\pi m kT}{h^2} \right)^{3/2} V.$$

That is, this integral does not include the ground state energy, but produce approximately the same answer when $T < T_c$. This means it actually represents the number of excited atoms in the system rather than the total number,

$$N_{\text{excited}} = 2.612 \left(\frac{2\pi mkT}{h^2} \right)^{3/2} V.$$

In conclusion, when the temperature is higher than T_c , the chemical potential is negative (compared to ϵ_0 , i.e. $\mu - \epsilon_0 < 0$) and almost all the atoms are in excited states, with the total number given by

$$N_{\text{excited}} = 2.612 \left(\frac{2\pi mkT_c}{h^2} \right)^{3/2} V, \quad [T > T_c].$$

At temperature smaller than T_c , the chemical potential is close to the ground state energy from negative and the number of excited atoms is

$$N_{\text{excited}} = 2.612 \left(\frac{2\pi mkT}{h^2} \right)^{3/2} V = \left(\frac{T}{T_c} \right)^{3/2} N, \quad [T < T_c].$$

The rest of the atom are in the ground state,

$$N_0 = \left[1 - \left(\frac{T}{T_c} \right)^{3/2} \right] N, \quad [T < T_c].$$

This effect of large fraction of bosons occupying the ground state abruptly below a critical temperature is called **Bose-Einstein condensation**. The critical temperature T_c is called the **condensation temperature**, and the ground-state atoms are called the **condensate**. You may already notice that condensation temperature is about the temperature at which the quantum volume $v_Q = (h^2/2\pi mkT)^{3/2} \sim V/N$. In other words, the condensates have their wavefunctions overlapping significantly. Note that for a given N , the value of kT_c scales like $N^{2/3}$ —it will be much greater than ϵ_0 when N is large. This gives the energy scales: $\epsilon_0 - \mu \ll \epsilon_0 \ll kT_c$ when $T < T_c$. For more interesting quantities for $T < T_c$, see Appendix A.3.4. Here are some results:

- Total energy:

$$U = 0.770 \left(\frac{T}{T_c} \right)^{3/2} NkT. \quad (7.30)$$

- Heat capacity:

$$C_V = 5.031 \left(\frac{2\pi mkT}{h^2} \right)^{3/2} V k \quad \text{or} \quad \frac{C_V}{Nk} = 1.926 \left(\frac{T}{T_c} \right)^{3/2}. \quad (7.31)$$

- Entropy:

$$S = 1.284 \left(\frac{T}{T_c} \right)^{3/2} Nk. \quad (7.32)$$

- Helmholtz free energy:

$$F = -0.514 \left(\frac{T}{T_c} \right)^{3/2} NkT. \quad (7.33)$$

- Pressure:

$$P = - \left(\frac{\partial F}{\partial V} \right)_{N,T} = 1.343 \left(\frac{2\pi mkT}{h^2} \right)^{3/2} kT. \quad (7.34)$$

Note that the pressure is independent of volume and the total number of particles. This is because all the integrals in A.3.4 is evaluated from 0 to ∞ . We are neglecting the $\epsilon = 0$ (the condensates) in the sum when turning to the integral. Since $V \propto T_c^{-3/2}$, as the volume decreases, the condensation temperature increases. If other quantities are fixed, when T_c becomes larger, the energy becomes smaller. The total particle number is not changing, so there must be more excited particles condensing to the ground state. Pressure is due to those energetic and excited particles. Reducing the volume of a Bose gas does not increase the pressure, unlike an ideal gas, because the number of excited particles is also decreasing.

The first achieved Bose-Einstein condensation is a gas of rubidium-87 in 1995. About 10^4 atoms were cooled down using laser cooling in a volume of $\sim 10^{-15} \text{ m}^3$. The condensation temperature is about 10^{-7} K . However, the confining potential is not a “box”, but a *anisotropic* harmonic oscillator potential. Appendix A.3.5 discusses an *isotropic* harmonic oscillator potential.

7.6.2 Why Does it Happen?

Why does Bose-Einstein condensation happens? We need to go all the way back to statistics and combinatorics.

Suppose there are N non-interacting *distinguishable* particles in a box. We can treat them using Boltzmann statistics, with each individual as a separate system. If there are many states that one particle can occupy and have the energy of order kT , then the partition function Z_1 would be quite large. The probability of finding the particle in the ground state, $1/Z_1$, is very small. This applies to each particle in the system, so few particles are in the ground state, and there is no Bose-Einstein condensation.

Let's analyze the same situation in another approach. Consider the system state instead of individual states. If the ground state energy is taken to be zero, then the Boltzmann factor of the system state with all particles being in the ground state is just 1. The system with total energy U has a Boltzmann factor $e^{-U/kT}$. If the particles have energy $\sim kT$, then $U \sim NkT$. The Boltzmann factor is suppressed like $e^{-NkT/kT} = e^{-N}$, hence the probability. Then how can the system not condensate into the ground state?

Think about the following example.

1. Take the ground state energy to be zero, and the energy spacing is equal in energy units. There are 5 non-interacting particles with one unit of energy in total. If the particles are distinguishable particles, we can choose any of the 5 particles to have that unit of energy. There are 5 possible combinations. If the particles are bosons, we can do the same, except that now the 5 combinations look the same because bosons are indistinguishable.
2. Now there are two units of energy. We can assign the particles to each energy level and see what will happen. Below shows the multiplicity and the combinations. (The energy level increases by one for each number to the right. For example, 32000 means three particles in the ground state, and two in the first excited state.) For distinguishable particles, we can choose two of them to occupy the first excited state, or one of them to occupy the second excited state:

$$32000, \quad \left[\Omega = \binom{5}{2} = 10 \right], \quad \text{and} \quad 40100, \quad \left[\Omega = \binom{5}{1} = 5 \right].$$

This gives the total multiplicity of $\Omega = 15$. For bosons, because they are indistinguishable, 32000 and 40100 stands for only two states, $\Omega = 2$.

3. Three units of energy. For distinguishable particles,

$$\begin{aligned} 23000 \quad \Omega &= \binom{5}{3} = 10, \\ 31100 \quad \Omega &= \frac{5!}{3!} = 20, \\ 40010 \quad \Omega &= \binom{5}{1} = 5. \end{aligned}$$

The total multiplicity is $\Omega = 35$. Again, bosons only have 23000, 31100, and 40010 but with no inner combinations, $\Omega = 3$.

This is just three energy units, but the difference in multiplicity of the two systems are already large. We can imagine that when the energy of the system is much higher, the distinguishable particle system has a much larger multiplicity than the bosonic system does. You may perceive this multiplicity as the *degeneracy*, n_s , of the system state. Since the probability is proportional to $n_s e^{-E(s)/kT}$, the distinguishable particle system has a much higher chance to be in excited states.

Indeed, the number of ways to arrange N distinguishable particles in Z_1 single-particle states is Z_1^N . Provided that $Z_1 \gg 1$, the suppression of e^{-N} is negligible. For a bosonic system, the number of arranging N indistinguishable particles among Z_1 single-particle states is mathematically the same as putting N units of energy in Z_1 oscillators in an Einstein solid.

$$\left(\begin{array}{c} \text{number of} \\ \text{system states} \end{array} \right) \sim \binom{N + Z_1 - 1}{N} \sim \begin{cases} (eZ_1/N)^N & Z_1 \gg N; \\ (eN/Z_1)^N & Z_1 \ll N. \end{cases}$$

When $Z_1 \gg N$, the factor of e^{-N} is negligible. This is the classical case. But when $Z_1 \ll N$ (this is equivalent to $V/N \ll v_Q$), e^{-N} dominates. For example, when $N = 100$ and $Z_1 = 25$, the Boltzmann factor is $\sim e^{-100} \approx 4 \times 10^{-44}$, but the number of states is only $\binom{124}{100} = 3 \times 10^{25}$.

In conclusion, identical particles, especially in quantum statistics, should be manipulated with care. This chapter is built on the fact that we know fermions or bosons of a given species are identical. The spin-statistics relations of fermions/bosons are theoretically proved in relativistic quantum mechanics. It is the indistinguishability forced us to count them in a different way from classical particles, and this is why quantum statistics often leads to surprises.

8 INTERACTING PARTICLES

The story of basic Thermal and Statistical Mechanics ends in the last chapter formally. This chapter is about systems with interacting particles, which are usually extremely hard to formulate, so we always want *ideal* systems. But in real world, nearly all systems are not ideal. In fact, this is why gases condense into liquids, materials acquire magnetization, etc. This forces us to make use of various approximation methods. There are two independent examples presented in this chapter, one is weakly interacting gases, and the other is the Ising model of ferromagnets.

8.1 Weakly Interacting Gases

We have seen nonideal gases modeled by the van der Waals equation. However, van der Waals equation only qualitatively predicts molecular interactions and phase transitions. We may want to see the behavior of nonideal gases by quantitative calculations from the first principle using tools of statistical mechanics. It is hard to model all kinds of interactions in gases, so in this section, we will only introduce *weakly* interacting gases.

8.1.1 The Partition Function

We will focus on the translational motion and put aside rotation or vibration. In Section 6.7, the state of a molecule is characterized by its energy and Z_{tr} is calculated by summing over all energies,

$$Z_1 = \sum_s e^{-\beta E_{\text{tr}}}$$

Alternatively, the state can also be characterized by a molecule's position and momentum,

$$Z_1 = \frac{1}{h^3} \int d^3r d^3p e^{-\beta E}.$$

There are in total six integrals. The integration includes all momentum space, and all positions within a box of volume V . A gas of N identical molecules will have a partition function

$$Z = \frac{1}{N!} \frac{1}{h^{3N}} \int d^3r_1 \cdots d^3r_N d^3p_1 \cdots d^3p_N e^{-\beta U}, \quad (8.1)$$

where U is the total energy of the system, and $1/N!$ counts for the overcounting of indistinguishable identical molecules. If the system were an ideal gas,

$$U_k = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \cdots + \frac{p_N^2}{2m}.$$

But we are considering a nonideal gas—there is potential energy between molecules. Let the total potential energy be U_p . The partition function can be written as

$$Z = \frac{1}{N!} \frac{1}{h^{3N}} \int d^3r_1 \cdots d^3r_N d^3p_1 \cdots d^3p_N e^{-\beta p_1^2/2m} \cdots e^{-\beta p_N^2/2m} e^{-\beta U_p}.$$

The momentum integrals are easy to evaluate because U_p is only a function of position. Each momentum integral contributes to a factor of

$$\int d^3p_i e^{-\beta p_i^2/2m} = (\sqrt{2\pi mkT})^3$$

There are N such integrals, so

$$Z = \frac{1}{N!} \left(\frac{2\pi mkT}{h^2} \right)^{3N/2} \int d^3r_1 \cdots d^3r_N e^{-\beta U_p} = \frac{Z_{\text{ideal}}}{V^N} \int d^3r_1 \cdots d^3r_N e^{-\beta U_p},$$

where Z_{ideal} is the ideal gas partition function (6.10). The remaining integral

$$Z_c = \frac{1}{V^N} \int d^3r_1 \cdots d^3r_N e^{-\beta U_p} \quad (8.2)$$

is called the **configuration integral**. The problem is to solve this integral.

8.1.2 The Cluster Expansion

We now make many assumptions to make calculation easier. First, assume that U_p can be written as a sum of potential energies between pairs of molecules:

$$U_p = u_{12} + u_{13} + \cdots + u_{1N} + u_{23} + \cdots + u_{N-1,N} = \sum_{\text{pairs}} u_{ij},$$

where u_{ij} is the potential energy between molecule i and molecule j . We also assume that u_{ij} depends only on the distance between the molecules, $|\mathbf{r}_i - \mathbf{r}_j|$. The configuration integral becomes

$$Z_c = \frac{1}{V^N} \int d^3r_1 \cdots d^3r_N \prod_{\text{pairs}} e^{-\beta u_{ij}}.$$

In reality, most potential goes to zero when the separation between two molecules is large. Assuming the gas is in the low density limit (interacting weakly), the potential energy satisfies $u_{ij} \ll kT$, so $e^{-\beta u_{ij}}$ is extremely close to 1. Let

$$e^{-\beta u_{ij}} = 1 + f_{ij}, \quad (8.3)$$

where f_{ij} is known as the **Mayors f -function**. The product of all Boltzmann factor is

$$\prod_{\text{pairs}} e^{-\beta u_{ij}} = \prod_{\text{pairs}} (1 + f_{ij}) = (1 + f_{12})(1 + f_{13}) \cdots (1 + f_{1N})(1 + f_{23}) \cdots (1 + f_{N-1,N}).$$

If we expand all these factors,

$$\prod_{\text{pairs}} e^{-\beta u_{ij}} = 1 + \sum_{\text{pairs}} f_{ij} + \sum_{\substack{\text{distinct} \\ \text{pairs}}} f_{ij} f_{kl} + \cdots,$$

and

$$Z_c = \frac{1}{V^N} \int d^3r_1 \cdots d^3r_N \left[1 + \sum_{\text{pairs}} f_{ij} + \sum_{\substack{\text{distinct} \\ \text{pairs}}} f_{ij} f_{kl} + \mathcal{O}(f^3) \right]. \quad (8.4)$$

The hope is that f_{ij} is small so that the terms involving more f -functions are less important. The first term in (8.4) yields

$$\frac{1}{V^N} \int d^3r_1 \cdots d^3r_N = \frac{1}{V^N} \left(\int d^3r \right)^N = 1.$$

Each r -integral is just the volume of the box. Terms with one f -function can reduce to only two integrals:

$$\frac{1}{V^N} \int d^3r_1 \cdots d^3r_N f_{ij} = \frac{1}{V^N} V^{N-2} \int d^3r_i d^3r_j f_{ij} = \frac{1}{V^2} \int d^3r_i d^3r_j f_{ij}.$$

This is because f_{ij} only depends on \mathbf{r}_i and \mathbf{r}_j . Since any two molecules can be 1 and 2, each pair f_{ij} produce the same result as f_{12} . Thus, the second term in (8.4) is

$$\frac{1}{V^N} \int d^3r_1 \cdots d^3r_N \sum_{\text{pairs}} f_{ij} = \frac{1}{2} \frac{N(N-1)}{V^2} \int d^3r_1 d^3r_2 f_{12}.$$

The factor $N(N-1)/2$ is the number of distinct pairs of f_{ij} . It is convenient to represent this expression (and other expressions later) by diagrams:

$$\bullet \text{---} \bullet = \frac{1}{2} \frac{N(N-1)}{V^2} \int d^3r_1 d^3r_2 f_{12}.$$

The rules that convert an integral to a diagram are the following:

1. Number the dots starting with 1, and each dot i corresponds to $(1/V) \int d^3r_i$. Multiply by N for the first dot, $N-1$ for the second, $N-2$ for the third, and so on.
2. A line connecting dots i and j corresponds to a factor of f_{ij} .

3. Divide by the symmetry factor of the diagram, which is the number of ways of numbering the dots without changing the corresponding product of f -functions. This is the number of permutations of dots that leave the diagram unchanged.

This type of diagrams also shows physical connections between molecules. We will see how this is in some examples below. The terms with two f -functions involve two cases: one is

$$\text{Diagram with 3 dots in a triangle} = \frac{1}{2} \frac{N(N-1)(N-2)}{V^3} \int d^3r_1 d^3r_2 d^3r_3 f_{12}f_{23}.$$

There is one molecule (labeled 2) interacting with other two molecules (labeled 1 and 3). The label of 1 and 3 can be interchanged, so there is a symmetry factor of 2. Other terms involving two f -functions are like

$$\left(\begin{array}{c} \bullet \\ \bullet \end{array} \begin{array}{c} \bullet \\ \bullet \end{array} \right) = \frac{1}{8} \frac{N(N-1)(N-2)(N-3)}{V^4} \int d^3r_1 d^3r_2 d^3r_3 d^3r_4 f_{12}f_{34}.$$

This represents two interactions between pairs of molecules. The symmetry factor is 8 because $f_{12} = f_{21}$, $f_{34} = f_{43}$, and (12) and (34) are interchangeable. Repeating such processes, the configuration integral can be written as a sum of diagrams:

$$Z_c = 1 + \text{Diagram with 2 dots} + \text{Diagram with 3 dots in a triangle} + \left(\begin{array}{c} \bullet \\ \bullet \end{array} \begin{array}{c} \bullet \\ \bullet \end{array} \right) + \text{Diagram with 4 dots in a square} + \text{Diagram with 4 dots in a chain} + \text{Diagram with 4 dots in a cycle} + \left(\begin{array}{c} \bullet \\ \bullet \end{array} \begin{array}{c} \bullet \\ \bullet \end{array} \begin{array}{c} \bullet \\ \bullet \end{array} \right) + \left(\begin{array}{c} \bullet \\ \bullet \\ \bullet \end{array} \right) + \dots$$

This is an example of a **diagrammatic perturbation series**. We can classify these terms by identical subdiagrams, such as

$$1 + \text{Diagram with 2 dots} + \left(\begin{array}{c} \bullet \\ \bullet \end{array} \begin{array}{c} \bullet \\ \bullet \end{array} \right) + \left(\begin{array}{c} \bullet \\ \bullet \\ \bullet \end{array} \right) + \dots$$

Taking the limit $N \gg 1$, $N = N - 1 = N - 2 = \dots$, it can be simplified to

$$1 + \text{Diagram with 2 dots} + \frac{1}{2} \left(\begin{array}{c} \bullet \\ \bullet \end{array} \right)^2 + \frac{1}{3!} \left(\begin{array}{c} \bullet \\ \bullet \end{array} \right)^3 + \dots = \exp \left(\begin{array}{c} \bullet \\ \bullet \end{array} \right).$$

With some mathematics beyond the scope of this notes, we can prove that in the large N limit,

$$Z_c = \exp \left(\text{Diagram with 2 dots} + \text{Diagram with 3 dots in a triangle} + \text{Diagram with 4 dots in a square} + \text{Diagram with 4 dots in a cycle} + \text{Diagram with 4 dots in a cube} + \dots \right).$$

These terms have a common point: they are connected diagrams, and are connected when any single dot are removed. Each term is called a **cluster**, and thus this formula for the configuration integral is called the **cluster expansion**.

Now with $Z = Z_{\text{ideal}}Z_c$, we are able to compute many thermodynamic quantities if we know how to evaluate the cluster expansion explicitly:

$$\begin{aligned} F &= -kT \ln Z = -kT \ln Z_{\text{ideal}} - kT \ln Z_c \\ &= -NkT \ln \left(\frac{V}{Nv_Q} \right) - kT \left(\text{Diagram with 2 dots} + \text{Diagram with 3 dots in a triangle} + \text{Diagram with 4 dots in a square} + \dots \right), \\ P &= - \left(\frac{\partial F}{\partial V} \right)_{N,T} = \frac{NkT}{V} + kT \frac{\partial}{\partial V} \left(\text{Diagram with 2 dots} + \text{Diagram with 3 dots in a triangle} + \text{Diagram with 4 dots in a square} + \dots \right), \quad \text{etc.} \end{aligned}$$

8.1.3 The Second Virial Coefficient

Consider the two-dot diagram:

$$\text{Diagram with 2 dots} = \frac{1}{2} \frac{N^2}{V^2} \int d^3r_1 d^3r_2 f_{12} = \frac{1}{2} \frac{N^2}{V^2} \int d^3r_1 \left[\int d^3r f(r) \right],$$

where $r \equiv |\mathbf{r}_2 - \mathbf{r}_1|$ and $f(r) = e^{-\beta u(r)} - 1$. We can say something about the r integral. It is independent of \mathbf{r}_1 and V . Here is the reasoning: assume that $f(r)$ goes to zero when r is a few times larger than the size of the molecule r_s and V is large. Then the r -integral only has some effective part from 0 to $r_{\text{eff}} = nr_s$, where n is some factor of order unity. No matter where \mathbf{r}_1 is, the value of r -integral is fixed, with the domain of integral within some sphere of radius r_{eff} . This breaks down only when \mathbf{r}_1 is near the wall of the box (the sphere is incomplete), but this chance is negligible if $V \gg r_s^3$. Therefore, the integral over \mathbf{r}_1 is simply a factor of V ,

$$\text{Diagram with 2 dots} = \frac{1}{2} \frac{N^2}{V} \int d^3r f(r).$$

Since the V dependence is clear, we can compute the pressure:

$$P = \frac{NkT}{V} + kT \frac{\partial}{\partial V} \left[\frac{1}{2} \frac{N^2}{V} \int d^3r f_{12}(r) \right] + \dots = \frac{NkT}{V} \left[1 - \frac{1}{2} \frac{N}{V} \int d^3r f(r) + \dots \right].$$

This is often written in the form of the **virial expansion**,

$$P = \frac{NkT}{V} \left[1 + \frac{B(T)}{(V/N)} + \frac{C(T)}{(V/N)^2} + \dots \right].$$

Our job is to compute the **second virial coefficient**:

$$B(T) = -\frac{1}{2} \int d^3r f(r). \quad (8.5)$$

We will use spherical coordinates, $d^3r = r^2 \sin \theta dr d\theta d\phi$, to evaluate $B(T)$,

$$B(T) = -2\pi \int_0^\infty r^2 f(r) dr = -2\pi \int_0^\infty r^2 \left[e^{-\beta u(r)} - 1 \right] dr.$$

Now it's time to specify $u(r)$. It should be a function that is weakly attractive at large distance and strongly repulsive at short distance. For neutral molecules with no permanent electric dipole moments, the long-distance force is $\sim 1/r^7$, resulting from spontaneous fluctuating dipole moment of a molecule. Hence this force give rise to a potential $\sim 1/r^6$. The exact form of repulsion potential is not crucial. For mathematical simplicity, we will use the **Lennard-Jones 6-12 potential**,

$$u(r) = u_0 \left[\left(\frac{r_0}{r} \right)^{12} - 2 \left(\frac{r_0}{r} \right)^6 \right], \quad (8.6)$$

where r_0 is the equilibrium position, and $u_0 = |u_{\min}|$ is the minimum depth of the potential well. Figure 8.1 shows the plot of $u(r)$ and $f(r)$ at different temperatures.

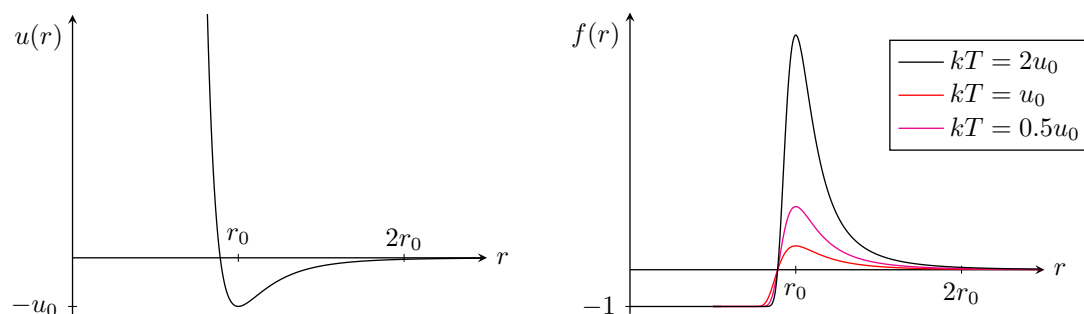


Figure 8.1: Left: the Lennard-Jones intermolecular potential function, with a strong repulsive region at small distances and weak attractive region at large distances. Right: the corresponding Mayer f -function, for three different temperatures.

Substituting $u(r)$ into $B(T)$ and evaluate the integral numerically (you can also inspect from the graph), we arrive at following results: At low temperatures, the integral of f is positive so the second virial coefficient is negative. The pressure is lower than that of an ideal gas. At high temperatures, the integral of f is negative so the second virial coefficient is positive. The pressure is higher than that of an ideal gas. For imple molecules, the shape of $B(T)$ predicted by the Lennard-Jones potential agrees very well with experiments.

8.2 The Ising Model of a Ferromagnet

In an ideal two-state paramagnet, it is assumed that magnetic dipoles only respond to external magnetic fields, but not their neighbors. Real world magnets have mutually interacting dipoles—some dipoles would prefer aligning parallel or anti-parallel—because dipoles themselves produce magnetic fields. A **ferromagnet** has its neighboring dipoles aligning parallel to each other (in the absence of external magnetic field). The most common example is iron, after which ferromagnet was named. An **anti-ferromagnet** has antiparallel neighboring dipoles. Some examples are Cr, NiO, and FeO.

At room temperature, a ferromagnet has a net nonzero magnetization. Raising the temperature causes random fluctuations that decrease the magnetization. Eventually, the net magnetization drops to zero when a ferromagnet is above a critical temperature

called the **Curie temperature**. Above that a ferromagnet acts like a paramagnet. For example, the Curie temperature of iron is 1043 K.

Below the Curie temperature, a ferromagnet can also have net zero magnetization. This is because it divides itself into small domains (but containing billions of dipoles). Each domain creates a net magnetic field, giving a tendency to magnetize neighboring domains in the opposite direction, and hence net zero magnetization. If the ferromagnet is heated up to a high enough temperature in an external field, all dipoles will then align with the field. Cooling it down to ordinary temperature “freezes” the dipoles, giving a *permanent* magnet.

In this section, we study a single domain in a ferromagnet and investigate the tendency of neighboring dipoles to align parallel to each other. There are two assumptions made: first ignore long-range magnetic interactions; second, assume that there is a preferred axis of magnetization that each dipole wants to align with. This is the **Ising model** of a ferromagnet. Let N be the total number of dipoles, and s_i be the state of the i th dipole. We will use $s_i = 1$ when the dipole is pointing up, and $s_i = -1$ when it is pointing down. When a pair of neighboring dipoles are parallel, the interaction energy is $+\epsilon$, while when they are anti-parallel, the interaction energy is $-\epsilon$. That is, if i and j are neighbors, their energy is $-\epsilon s_i s_j$. The total energy of the whole system (the domain) is

$$U = -\epsilon \sum_{\text{neighboring pairs } i,j} s_i s_j.$$

The partition function is

$$Z = \sum_{\{s_i\}} e^{-\beta U},$$

where the sum is over all possible arrangements of dipole alignments. If there are N two-state dipoles, then the number of terms in the sum is 2^N .

8.2.1 Exact Solution in One Dimension

Suppose now all dipoles are aligning in a straight line. Each dipole (ignore the edges) then have two nearest neighbors. A possible configuration could be:

$$\begin{array}{ccccccccc} \uparrow & \downarrow & \uparrow & \downarrow & \downarrow & \cdots & \uparrow \\ i = 1 & 2 & 3 & 4 & 5 & \cdots & N \\ s_i = 1 & -1 & 1 & -1 & -1 & \cdots & 1 \end{array}$$

In the absence of an external field, the total energy and the partition function of this one-dimensional Ising model are

$$U = -\epsilon(s_1 s_2 + s_2 s_3 + s_3 s_4 + \cdots + s_{N-1} s_N).$$

$$Z = \sum_{s_1} \sum_{s_2} \cdots \sum_{s_N} e^{\beta \epsilon s_1 s_2} e^{\beta \epsilon s_2 s_3} \cdots e^{\beta \epsilon s_{N-1} s_N},$$

where each s_i can be -1 and $+1$. The final sum follows (regardless of the sign of s_{N-1})

$$\sum_{s_N} e^{\beta \epsilon s_{N-1} s_N} = e^{\beta \epsilon} + e^{-\beta \epsilon} = 2 \cosh \beta \epsilon.$$

Since this is now independent of s_{N-1} , the next sum over s_{N-1} have the same result,

$$\sum_{s_{N-1}} \sum_{s_N} e^{\beta \epsilon s_{N-2} s_{N-1}} e^{\beta \epsilon s_{N-1} s_N} = \sum_{s_{N-1}} e^{\beta \epsilon s_{N-2} s_{N-1}} (2 \cosh \beta \epsilon) = (2 \cosh \beta \epsilon)^2.$$

Keep doing the sum in the same way, there are $N - 1$ factors of $2 \cosh \beta \epsilon$ up to s_1 :

$$Z = \sum_{s_1} (2 \cosh \beta \epsilon)^{N-1} = 2^N (\cosh \beta \epsilon)^{N-1} \approx (2 \cosh \beta \epsilon)^N.$$

The last sum over s_1 does not have an exponent. The approximation is valid when N is large. As usual, the first thing to calculate is the average energy of the system:

$$\langle U \rangle = -\frac{\partial}{\partial \beta} \ln Z = -N \epsilon \tanh \beta \epsilon.$$

The average energy goes to $-N\epsilon$ as $T \rightarrow 0$, which is the least possible energy where all dipoles are parallel. As $T \rightarrow \infty$, the average energy goes to zero. This means at high temperature, the dipoles are randomly aligned, with half parallel and half anti-parallel in pairs. Note that this result resembles the Ising paramagnet if you replace ϵ with μB . There is no abrupt transition of magnetization or average energy of this one-dimensional model because \tanh is smooth. However, a three-dimensional ferromagnet does have transitions at some critical temperature. We should look at higher dimensions too see how this works.

In 1940s, the *two*-dimensional Ising model of a square lattice was solve by Lars Onsager. Onsager obtained the exact partition function as $N \rightarrow \infty$, and showed that there is a critical temperature. The solution is extremely difficult mathematically. There is no exact solution to three-dimensional Ising model found.

8.2.2 The Mean Field Approximation

In this small section we will present an important approximation to the Ising model and see why dimensionality matters. For a single dipole, the number of nearest neighbors n of it is

$$n = \begin{cases} 2 & \text{in one dimension,} \\ 4 & \text{in two dimensions (square lattice),} \\ 6 & \text{in three dimensions (simple cubic lattice),} \\ 8 & \text{in three dimensions (body-centered cubic lattice),} \\ 12 & \text{in three dimensions (face-centered cubic lattice).} \end{cases}$$

The energy of this dipole due to its neighbors is simply

$$E_{\uparrow} = -\epsilon \sum_{\text{neighbors}} s_{\text{neighbor}} = -\epsilon n \langle s \rangle \quad \text{and} \quad E_{\downarrow} = +\epsilon n \langle s \rangle.$$

The partition function for this single dipole is

$$Z_i = e^{\beta \epsilon n \langle s \rangle} + e^{-\beta \epsilon n \langle s \rangle} = 2 \cosh(\beta \epsilon n \langle s \rangle).$$

This gives the expectation value of its spin alignment:

$$\langle s_i \rangle = \frac{1}{Z_i} \left[1 \cdot e^{\beta \epsilon n \langle s \rangle} + (-1) \cdot e^{-\beta \epsilon n \langle s \rangle} \right] = \frac{2 \sinh(\beta \epsilon n \langle s \rangle)}{2 \cosh(\beta \epsilon n \langle s \rangle)} = \tanh(\beta \epsilon n \langle s \rangle).$$

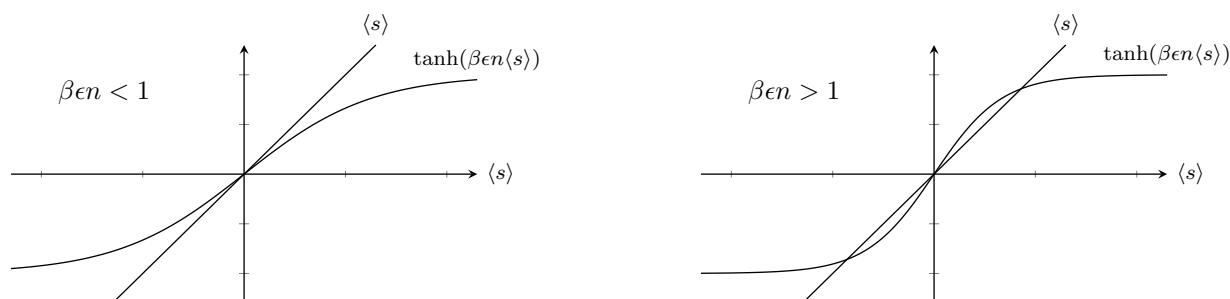
The LHS of this equation is the average value of spin alignment of any dipole (again neglect edges). The RHS contains $\langle s \rangle$, which is the average instantaneous alignments of the neighbors of this dipole. The **mean field approximation** pretends that the two average values are the same,

$$\langle s_i \rangle = \langle s \rangle.$$

In the mean field approximation, the equation above produces an equation of $\langle s \rangle$.

$$\langle s \rangle = \tanh(\beta \epsilon n \langle s \rangle).$$

Since this is a transcendental equation, we will look at graphical solutions:



The number of solutions of this equation depends on $\beta \epsilon n$. When $\beta \epsilon n < 1$ or $kT > n\epsilon$, the slope of \tanh at $\langle s \rangle = 0$ is less than one.

There is only one solution at $\langle s \rangle = 0$ in this case. It is a stable solution because when a thermal fluctuation increase $\langle s \rangle$, $\tanh(\beta\epsilon n\langle s \rangle)$ tells where $\langle s \rangle$ should be. Since \tanh is less than the current value of $\langle s \rangle$, $\langle s \rangle$ tends to decrease back to zero. When $\beta\epsilon n > 1$ or $kT < n\epsilon$, there will be three solutions. By the same reasoning above, we can argue that at $\langle s \rangle = 0$, this solution is unstable. Any fluctuation will bifurcate it into two stable solutions, one positive and another negative. If the fluctuation is random, the system is equally likely to have positive or negative magnetization.

The critical temperature is then $kT_c = n\epsilon$. Below T_c , the system becomes magnetized. As discussed in the one-dimensional case, there is no abrupt transition of magnetization. Thus, the mean field approximation is not a good approximation in one dimension, but the accuracy improves with dimensionality.

Acknowledgements: I thank Professor David Weld, my instructor of the course Phys 119A: Thermal and Statistical Mechanics.

Reference: Schroeder, Daniel V. *An Introduction to Thermal Physics*. Addison Wesley, 2000.

A PROOFS AND DERIVATIONS

A.1 Energy and Entropy

A.1.1 Negative Heat Capacities

(a) Consider a system of just two particles of identical mass m , orbiting in circles of radius r about their center of mass (CM). The separation between the two particles are then $2r$. The two particles are held by a centripetal force (so that they are in circular motion) and such force is gravity:

$$F_c = F_g.$$

By the centripetal force formula $F_c = mv^2/r$, where v is the velocity of each particle relative to CM, and the law of universal gravitation,

$$\frac{mv^2}{r} = \frac{Gm^2}{(2r)^2}.$$

Rearranging some terms:

$$2mv^2 = \frac{Gm^2}{2r}.$$

The LHS of the equation is twice the total kinetic energy of the particles, where the total kinetic energy of the particles is $U_k = 2(\frac{1}{2}mv^2) = mv^2$. The RHS is the negative of the gravitational potential energy of the system, $U_p = -Gm^2/2r$. Hence, we arrive at

$$-2U_k = U_p,$$

which says that the gravitational potential energy is -2 times the total kinetic energy of the system.

(b) The conclusion of part (a) turns out to be true on average for any system of particles held together by mutual gravitational attraction,

$$\bar{U}_p = -2\bar{U}_k.$$

This result is known as the **virial theorem**. Now we compute the total internal energy of the system,

$$U = U_p + U_k = -2U_k + U_k = -U_k.$$

This means if we add some energy to the system, the average total kinetic energy should decrease.

(c) A star can be modeled as a gas of particles that interact with each other only gravitationally. According to the equipartition theorem, the average kinetic energy of the particles in such a star should be $\frac{3}{2}kT$, where T is the average temperature. The total energy of a star is

$$U = -U_k = -\frac{3}{2}NkT,$$

where N is the number of particles in the star. The heat capacity is

$$C = \frac{\partial U}{\partial T} = \boxed{-\frac{3}{2}Nk}.$$

Example A.1. Temperature of the Sun

A star of mass M and radius R should have a total potential energy of $U_p \sim -GM^2/R$ (there is some factor of order 1). Assume that the sun is uniform, and it is made entirely of protons and electrons (which is reasonable because the sun is mostly composed of hydrogen). Given that the solar mass is 2×10^{30} kg and its radius is 7×10^8 m, we want to estimate the temperature of the

sum. First we relate temperature to mass and radius, so we invoke the virial theorem:

$$U_p = -2U_k \implies -\frac{GM^2}{R} = -2\left(\frac{3}{2}NkT\right),$$

which gives

$$T = -\frac{1}{3} \frac{GM^2}{NkR}.$$

The sun is about neutral, so $N_{\text{proton}} = N_{\text{electron}}$, or

$$N = 2N_{\text{proton}}.$$

The proton is much heavier than the electron. Thus approximately the sun's mass is all contributed by protons. We can now determine N ,

$$N = 2N_{\text{proton}} = \frac{2M}{m_p}.$$

Finally,

$$T = -\frac{1}{3} \frac{GM^2}{NkR} = -\frac{1}{6} \frac{GMm_p}{kR} \approx 3.8 \times 10^6 \text{ K}.$$

This temperature is much higher than the surface of the sun, but about 4 times lower than the central temperature.

A.1.2 Multiplicity of a Low-Temperature Einstein Solid

To get an expression for the multiplicity of an Einstein solid in low temperature limit ($q \ll N$), we will repeat the method for high temperature limit. First, take the logarithm of the multiplicity of an Einstein solid (2.1) using Stirling's approximation,

$$\ln \Omega \approx (N + q) \ln(N + q) - q \ln q - N \ln N.$$

The first logarithm can be approximated using $q \ll N$,

$$\ln(N + q) = \ln \left[N \left(1 + \frac{q}{N} \right) \right] = \ln N + \ln \left(1 + \frac{q}{N} \right) \approx \ln N + \frac{q}{N}.$$

Thus, the multiplicity becomes

$$\ln \Omega \approx q \ln N + q + q \left(\frac{q}{N} \right) - q \ln q \approx q \ln \frac{N}{q} + q,$$

where q^2/N is dropped because it is small compared to other terms. Re-exponentiating to get the multiplicity for a low-temperature Einstein solid,

$$\Omega \approx \left(\frac{eN}{q} \right)^q.$$

A.1.3 Entropy of a Black Hole

Throwing things into a black hole is an irreversible process, both mechanically and thermodynamically: adding mass to a black hole increases its entropy. (a) By dimensional analysis, the radius of a black hole with mass M is

$$r \sim \frac{GM}{c^2}.$$

(b) For a large system, the entropy is often at the order of $\sim Nk$. Thus, entropy can be seen as a measure of particle number in a system. The black hole does not care about what its mass is composed of. To maximize entropy, the black hole should contain as many particles as possible. Since energy is related to mass by $E = Mc^2$, the particles should have the lowest possible energy to maximize particle number.

(c) The particle with lowest possible energy can be a long-wavelength photon. But the wavelength cannot be larger than the size of the black hole. Thus, the energy of the photon is

$$E_\gamma = \frac{hc}{\lambda} \quad \text{with} \quad \lambda \lesssim \frac{GM}{c^2}.$$

The entropy of a black hole is to the order of

$$S \sim \frac{Mc^2}{E_\gamma} k = \frac{Mc^2}{hc/\lambda} k = \frac{GM^2}{hc} k.$$

Example A.2. One-solar-mass black hole

A one-solar-mass black hole has an entropy $\sim 10^{77}k$. For comparison, the sun has entropy about $10^{57}k$ because there are about $\sim 10^{57}$ particles in it. You will need 10^{20} ordinary stars to equal the entropy of a solar-mass black hole. Note that the entropy of a black hole is proportional to M^2 , a supermassive black hole would have enormous entropy.

A.1.4 Entropy of an Ising Paramagnet

N_\uparrow and N_\downarrow are related by energy U by (3.3),

$$N_\uparrow = \frac{1}{2} \left(N - \frac{U}{\mu B} \right), \quad \text{and} \quad N_\downarrow = \frac{1}{2} \left(N + \frac{U}{\mu B} \right).$$

Once we get U in terms of T (3.7),

$$\frac{U}{\mu B} = -N \tanh x, \quad \text{where} \quad x = \mu B/kT.$$

This gives N_\downarrow and N_\uparrow as a function of x ,

$$N_\uparrow = \frac{N}{2}(1 + \tanh x) = \frac{N}{2} \left(\frac{\cosh x + \sinh x}{\cosh x} \right) = \frac{Ne^x}{2 \cosh x},$$

$$N_\downarrow = \frac{N}{2}(1 - \tanh x) = \frac{N}{2} \left(\frac{\cosh x - \sinh x}{\cosh x} \right) = \frac{Ne^{-x}}{2 \cosh x}.$$

In terms of N_\uparrow and N_\downarrow , the entropy is

$$S/k = N \ln N - N_\uparrow \ln N_\uparrow - N_\downarrow \ln N_\downarrow.$$

Plug in N_\uparrow and N_\downarrow ,

$$\begin{aligned} \frac{S}{Nk} &= \ln N - \frac{e^x}{2 \cosh x} \ln \left(\frac{Ne^x}{2 \cosh x} \right) - \frac{e^{-x}}{2 \cosh x} \ln \left(\frac{Ne^{-x}}{2 \cosh x} \right) \\ &= \ln N - \frac{e^x}{2 \cosh x} [\ln N + x - \ln(2 \cosh x)] - \frac{e^{-x}}{2 \cosh x} [\ln N - x - \ln(2 \cosh x)] \\ &= -x \left(\frac{e^x - e^{-x}}{2 \cosh x} \right) + \left(\frac{e^x + e^{-x}}{2 \cosh x} \right) \ln(2 \cosh x) \\ &= \ln(2 \cosh x) - x \tanh x, \end{aligned}$$

or

$$S = Nk[\ln(2 \cosh x) - x \tanh x].$$

A.2 Boltzmann Statistics

A.2.1 Energy of Quantum Harmonic Oscillators

It is known that quantum harmonic oscillators have equal spacing of energy levels. If we let the ground state energy to be $E_0 = 0$, then $E_n = n\hbar\omega$. There is a useful Taylor series that converges for $|x| < 1$,

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots$$

The partition function of an harmonic oscillator is

$$Z = \sum_{n=0}^{\infty} e^{-\beta n \hbar \omega} = \sum_{n=0}^{\infty} (e^{-\beta \hbar \omega})^n = \frac{1}{1 - e^{-\beta \hbar \omega}}.$$

The average energy is

$$\langle E \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -(1 - e^{-\beta \hbar \omega}) \frac{\partial}{\partial \beta} \frac{1}{1 - e^{-\beta \hbar \omega}} = \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1}.$$

This gives the total energy of an N -harmonic oscillator system at temperature T ,

$$U = N \langle E \rangle = \frac{N \hbar \omega}{e^{\hbar \omega / kT} - 1}.$$

From this we can get the heat capacity,

$$C = \frac{\partial U}{\partial T} = \frac{\partial}{\partial T} \frac{N \hbar \omega}{e^{\hbar \omega / kT} - 1} = \frac{N \hbar^2 \omega^2}{kT^2} \frac{e^{\hbar \omega / kT}}{(e^{\hbar \omega / kT} - 1)^2}.$$

In the limit $T \rightarrow 0$, $C \rightarrow 0$. At large T , $\hbar \omega / kT \ll 1$, so we can Taylor expand $e^{\hbar \omega / kT} \approx 1 + \hbar \omega / kT$ for the denominator to cancel the 1 and $e^{\hbar \omega / kT} \approx 1$ for the numerator:

$$C \approx \frac{N \hbar^2 \omega^2}{kT^2} \frac{1}{(\hbar \omega / kT)^2} = Nk,$$

as predicted by equipartition theorem.

A.2.2 Linear Degrees of Freedom

Consider a classical “degree of freedom” that is linear rather than quadratic: $E = c|q|$ for some constant c . The average energy is $\langle E \rangle = kT$.

Proof. The partition function of a particle with energy $E = c|q|$ is

$$\begin{aligned} Z &= \sum_q e^{-\beta c|q|} = \frac{1}{\delta q} \sum_q e^{-\beta c|q|} \delta q \\ &\approx \frac{1}{\delta q} \int_{-\infty}^{+\infty} e^{-\beta c|q|} dq \\ &= \frac{2}{\delta q} \frac{1}{\beta c} = c' \frac{1}{\beta}, \end{aligned}$$

where $c' = 2/(c \delta q)$. The average energy is given by

$$\langle E \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\beta}{c'} \frac{\partial}{\partial \beta} \left(\frac{c'}{\beta} \right) = \frac{1}{\beta} = kT.$$

□

A.3 Quantum Statistics

A.3.1 Sommerfeld Expansion

Chemical Potential

The integral for N is

$$N = \int_0^\infty g(\epsilon) \bar{n}_{\text{FD}}(\epsilon) d\epsilon = g_0 \int_0^\infty \epsilon^{1/2} \bar{n}_{\text{FD}} d\epsilon,$$

where g_0 is some constants before $\sqrt{\epsilon}$ in $g(\epsilon)$. The interesting region is near $\epsilon \approx \mu$ (see Figure 7.2), so we should isolate this region by integration by parts:

$$N = \frac{2}{3} g_0 \epsilon^{3/2} \bar{n}_{\text{FD}} \Big|_0^\infty + \frac{2}{3} g_0 \int_0^\infty \epsilon^{3/2} \left(-\frac{d\bar{n}_{\text{FD}}}{d\epsilon} \right) d\epsilon.$$

By making a change of variable $x = (\epsilon - \mu)/kT$, the derivative $d\bar{n}_{\text{FD}}/d\epsilon$ is

$$-\frac{d\bar{n}_{\text{FD}}}{d\epsilon} = -\frac{d}{d\epsilon} \left[e^{(\epsilon - \mu)/kT} + 1 \right]^{-1} = \frac{1}{kT} \frac{e^x}{(e^x + 1)^2}.$$

The integral becomes

$$N = \frac{2}{3}g_0 \int_0^\infty \frac{1}{kT} \frac{e^x}{(e^x + 1)^2} \epsilon^{3/2} d\epsilon = \frac{2}{3}g_0 \int_{-\mu/kT}^\infty \frac{e^x}{(e^x + 1)^2} \epsilon^{3/2} dx.$$

The integrand decays exponentially when $|\epsilon - \mu| \gg kT$. Now make two approximations

1. Extend the lower limit of the integrand to $-\infty$ to make the integral symmetric.
2. Take the first few terms of the Taylor series about $\epsilon = \mu$.

$$\epsilon^{3/2} = \mu^{3/2} + \frac{3}{2}(\epsilon - \mu)\mu^{1/2} + \frac{3}{8}(\epsilon - \mu)^2\mu^{-1/2} + \dots$$

The integral becomes integrable with just polynomials and exponential,

$$N = \frac{2}{3}g_0 \int_{-\infty}^{+\infty} \frac{e^x}{(e^x + 1)^2} \left[\mu^{3/2} + \frac{3}{2}xkT\mu^{1/2} + \frac{3}{8}(xkT)^2\mu^{-1/2} + \dots \right] dx.$$

The first term:

$$\int_{-\infty}^{+\infty} \frac{e^x}{(e^x + 1)^2} dx = \int_{-\infty}^{+\infty} -\frac{d\bar{n}_{\text{FD}}}{d\epsilon} d\epsilon = \bar{n}_{\text{FD}}(-\infty) - \bar{n}_{\text{FD}}(\infty) = 1.$$

The second term is an odd function of x , so

$$\int_{-\infty}^{+\infty} \frac{xe^x}{(e^x + 1)^2} dx = 0.$$

The third integral can be evaluated analytically. The result is

$$\int_{-\infty}^{+\infty} \frac{x^2 e^x}{(e^x + 1)^2} dx = \frac{\pi^2}{3}.$$

Summing up these results give the total number of electrons:

$$N = \frac{2}{3}g_0\mu^{3/2} + \frac{1}{4}g_0(kT)^2\mu^{-1/2} \cdot \frac{\pi^2}{3} + \dots = N \left[N \left(\frac{\mu}{\epsilon_F} \right)^{3/2} + \frac{\pi^2}{8} \frac{(kT)^2}{\epsilon_F^{3/2} \mu^{1/2}} + \dots \right]$$

Canceling the N , we see that $\mu/\epsilon_F \approx 1$ because of the low temperature limit ($kT/\epsilon_F \ll 1$). The correction term is small, so we can approximate $\mu \approx \epsilon_F$. The equation for μ/ϵ_F becomes

$$1 = \left(\frac{\mu}{\epsilon_F} \right)^{3/2} + \frac{\pi^2}{8} \frac{(kT)^2}{\epsilon_F^2} + \dots$$

Solve for μ/ϵ_F :

$$\frac{\mu}{\epsilon_F} = \left[1 - \frac{\pi^2}{8} \left(\frac{kT}{\epsilon_F} \right)^2 + \dots \right]^{2/3} = 1 - \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 + \dots$$

Energy

The U -integral is

$$U = \int_0^\infty \epsilon g(\epsilon) \bar{n}_{\text{FD}}(\epsilon) d\epsilon = g_0 \int_0^\infty \epsilon^{3/2} \bar{n}_{\text{FD}}(\epsilon) d\epsilon.$$

Integrating by parts,

$$U = \cancel{\frac{2}{5}g_0\epsilon^{5/2}\bar{n}_{\text{FD}}(\epsilon)} \Big|_0^\infty + \frac{2}{5}g_0 \int_0^\infty \epsilon^{5/2} \left(-\frac{d\bar{n}_{\text{FD}}}{d\epsilon} \right) d\epsilon.$$

Let $x = (\epsilon - \mu)/kT$,

$$U = \frac{2}{5}g_0 \int_0^\infty \frac{1}{kT} \frac{e^x}{(e^x + 1)^2} \epsilon^{5/2} d\epsilon = \frac{2}{5}g_0 \int_{-\mu/kT}^\infty \frac{e^x}{(e^x + 1)^2} \epsilon^{5/2} dx.$$

Do the same approximation as in the chemical potential case: extend the lower limit of the integral to $-\infty$ and Taylor expand $\epsilon^{5/2}$,

$$\epsilon^{5/2} = \mu^{5/2} + \frac{5}{2}(\epsilon - \mu)\mu^{3/2} + \frac{15}{8}(\epsilon - \mu)^2\mu^{1/2} + \dots = \mu^{5/2} + \frac{5}{2}(xkT)\mu^{3/2} + \frac{15}{8}(xkT)^2\mu^{1/2} + \dots$$

The three integrals are the same: the first one is 1; the second one is zero; the third one is $\pi/3$, so

$$\begin{aligned} U &\approx \frac{2}{5}g_0\mu^{5/2} + \frac{3}{4}g_0(kT)^2\mu^{1/2} \cdot \frac{\pi^2}{3} + \dots \\ &= \frac{2}{5}g_0\mu^{5/2} + \frac{\pi^2}{4}g_0(kT)^2\mu^{1/2} + \dots \\ &= \frac{3}{5}\frac{N\mu^{5/2}}{\epsilon_F^{3/2}} + \frac{3\pi^2}{8}\frac{N(kT)^2}{\epsilon_F} + \dots \end{aligned}$$

where in the second term $\mu = \epsilon_F$ is used. Now we need to plug in μ ,

$$\mu^{5/2} = \epsilon_F^{5/2} \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 + \dots \right]^{5/2} = \epsilon_F^{5/2} \left[1 - \frac{5\pi^2}{24} \left(\frac{kT}{\epsilon_F} \right)^2 + \dots \right].$$

Finally,

$$U \approx \frac{3}{5}N\epsilon_F \left[1 - \frac{5\pi^2}{24} \left(\frac{kT}{\epsilon_F} \right)^2 \right] + \frac{3\pi^2}{8}\frac{N(kT)^2}{\epsilon_F} = \frac{3}{5}N\epsilon_F + \frac{\pi^2}{4}\frac{N(kT)^2}{\epsilon_F}.$$

A.3.2 Photon Gas Helmholtz Free Energy

We already know the partition function of one photon/mode,

$$Z = \frac{1}{1 - e^{-\epsilon/kT}}.$$

Thus the Helmholtz free energy of this mode is

$$F = -kT \ln Z = kT \ln(1 - e^{-\epsilon/kT}).$$

Summing over all modes,

$$F = 2 \sum_{n_x, n_y, n_z} kT \ln(1 - e^{-\epsilon/kT}) = 2kT \cdot \frac{\pi}{2} \int_0^\infty n^2 \ln(1 - e^{-\epsilon/kT}) dn.$$

As usual, we sum over the first octant in the n -space, and the factor of 2 accounts for two polarizations of photons. By the substitution $x = \epsilon/kT = hc n/2LkT$,

$$F = \pi kT \left(\frac{2LkT}{hc} \right)^3 \int_0^\infty x^2 \ln(1 - e^{-x}) dx = 8\pi V \frac{(kT)^4}{(hc)^3} \int_0^\infty x^2 \ln(1 - e^{-x}) dx.$$

Integrating by parts,

$$F = 8\pi V \frac{(kT)^4}{(hc)^3} \left[\frac{x^3}{3} \ln(1 - e^{-x}) \Big|_0^\infty - \int_0^\infty \frac{x^3}{3} \frac{e^{-x}}{1 - e^{-x}} dx \right] = -\frac{8\pi V}{3} \int_0^\infty \frac{x^3}{e^x - 1} dx = -\frac{1}{3}U.$$

A.3.3 Photon Spectrum in Terms of Wavelength

The energy density per unit photon energy $u(\epsilon)$ and the total energy density U/V are given by

$$u(\epsilon) = \frac{8\pi}{(hc)^3} \frac{\epsilon^3}{e^{\epsilon/kT} - 1}, \quad \frac{U}{V} = \frac{8\pi}{(hc)^3} \int_0^\infty \frac{\epsilon^3}{e^{\epsilon/kT} - 1} d\epsilon.$$

Making the change of variable $\epsilon = hc/\lambda$ and $d\epsilon = -(hc/\lambda^2) d\lambda$, the total energy density should not change, so

$$\frac{U}{V} = \frac{8\pi}{(hc)^3} \int_0^\infty \frac{\epsilon^3 d\epsilon}{e^{\epsilon/kT} - 1} = -\frac{8\pi}{(hc)^3} (hc)^4 \int_\infty^0 \frac{d\lambda/\lambda^5}{e^{hc/kT\lambda} - 1} = 8\pi hc \int_0^\infty \frac{d\lambda/\lambda^5}{e^{hc/kT\lambda} - 1}.$$

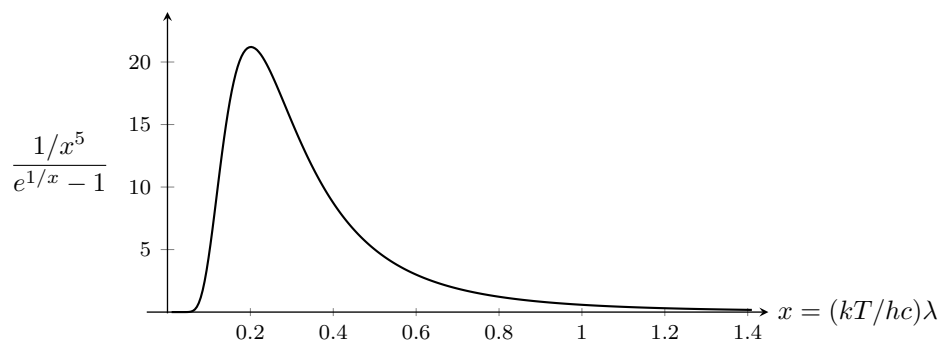
The integrand is the photon spectrum per unit wavelength:

$$u(\lambda) = \frac{8\pi hc}{\lambda^5} \frac{1}{e^{hc/kT\lambda} - 1}.$$

Making another change of variable $x = (kT/hc)\lambda$,

$$\frac{U}{V} = \frac{8\pi(kT)^4}{(hc)^3} \int_0^\infty \frac{1/x^5}{e^{1/x} - 1} dx.$$

The following is a plot of this integrand



A.3.4 Below the Condensation Temperature

The total energy of a Bose gas at a certain temperature T is the sum

$$U = \sum_s \frac{\epsilon}{e^{(\epsilon_s - \mu)/kT} - 1} \approx \int_0^\infty \epsilon \frac{g(\epsilon)}{e^{(\epsilon - \mu)/kT} - 1} d\epsilon = \frac{2}{\sqrt{\pi}} \left(\frac{2\pi m}{h^2} \right)^{3/2} V \int_0^\infty \frac{\epsilon^{3/2}}{e^{(\epsilon - \mu)/kT} - 1} d\epsilon.$$

For $T < T_c$, we can just set $\mu = 0$ because it is so close to $\epsilon_0 = 0$. Then make a change of variable $x = \epsilon/kT$ to evaluate the integral:

$$U = \frac{2}{\sqrt{\pi}} \left(\frac{2\pi m}{h^2} \right)^{3/2} V (kT)^{5/2} \int_0^\infty \frac{x^{3/2}}{e^x - 1} dx = \frac{2}{\sqrt{\pi}} \left(\frac{2\pi m}{h^2} \right)^{3/2} V (kT)^{5/2} \cdot 1.783.$$

This integral is done numerically. The heat capacity is given by

$$C_V = \frac{\partial U}{\partial T} = \frac{5}{2} \cdot 1.783 \cdot \frac{2}{\sqrt{\pi}} \left(\frac{2\pi m}{h^2} \right)^{3/2} V (kT)^{3/2} k = 5.031 \left(\frac{2\pi m k T}{h^2} \right)^{3/2} V k.$$

Both the heat capacity and total energy can be expressed in terms of N and T_c , using

$$N = 2.612 \left(\frac{2\pi m k T_c}{h^2} \right)^{3/2} V,$$

so

$$C_V = \frac{C_V}{Nk} = \frac{5.031}{2.612} \left(\frac{T}{T_c} \right)^{3/2} = 1.926 \left(\frac{T}{T_c} \right)^{3/2}, \quad U = \frac{2}{\sqrt{\pi}} \frac{1.783}{2.612} \left(\frac{T}{T_c} \right)^{3/2} NkT = 0.770 \left(\frac{T}{T_c} \right)^{3/2} NkT.$$

Once the heat capacity is known all the way to $T = 0$, the entropy of a Bose gas can be determined,

$$S = \int_0^T \frac{C_V}{T} dT = \frac{(1.926 Nk)}{T_c^{3/2}} \int_0^T T^{1/2} dT = \frac{1.926 Nk}{T_c^{3/2}} \cdot \frac{2}{3} T^{3/2}.$$

Thus, for any $T < T_c$,

$$S = 1.284 \left(\frac{T}{T_c} \right)^{3/2} Nk.$$

Combining the total energy, temperature, and entropy gives the Helmholtz free energy,

$$F = U - TS = -0.514 \left(\frac{T}{T_c} \right)^{3/2} NkT.$$

To get the pressure from $P = -(\partial F / \partial V)_{N,T}$, we need to express the Helmholtz free energy in terms of pressure. The only way to do that is to eliminate T_c ,

$$T_c^{-3/2} = 2.612 \frac{V}{N} \left(\frac{2\pi m k}{h^2} \right)^{3/2}.$$

Plug this into F ,

$$F = -0.514 \cdot 2.612 \frac{V}{N} \left(\frac{2\pi mkT}{h^2} \right)^{3/2} NkT = 1.343 \frac{V}{N} \left(\frac{2\pi mkT}{h^2} \right)^{3/2} NkT.$$

Finally, the pressure is

$$P = - \left(\frac{\partial F}{\partial V} \right)_{N,T} = 1.343 \left(\frac{2\pi mkT}{h^2} \right)^{3/2} kT.$$

A.3.5 Bose Gas in a Harmonic Oscillator Potential

Consider a gas of N identical spin-0 bosons confined by an isotropic three-dimensional harmonic oscillator potential, $V = \frac{1}{2}m\omega^2 r^2$. The energy levels in this potential are $\epsilon = n\hbar\omega = nh\nu$, where n is any nonnegative integer. The degeneracy of level n is $(n+1)(n+2)/2$. If we assume $n \gg 1$, the degeneracy of level n can be approximated as $n^2/2$. The total number of states below $n_{\max} = \epsilon_{\max}/h\nu$ is

$$N_{\text{states}} = \int_0^{n_{\max}} \frac{n^2}{2} dn = \int_0^{\epsilon_{\max}} \frac{\epsilon^2}{2(h\nu)^2} \frac{d\epsilon}{h\nu} = \int_0^{\epsilon_{\max}} \frac{\epsilon^2}{2(h\nu)^3} d\epsilon,$$

where we have used $n = \epsilon/h\nu$ and $dn = d\epsilon/h\nu$. The density of states is the integrand

$$g(\epsilon) = \frac{\epsilon^2}{2(h\nu)^3}.$$

With the chemical potential set to zero (just like the guess for particle in a box case), the total number of particles is the integral

$$N = \int_0^{\infty} g(\epsilon) \frac{1}{e^{\epsilon/kT} - 1} d\epsilon = \frac{1}{2(h\nu)^3} \int_0^{\infty} \frac{\epsilon^2}{e^{\epsilon/kT} - 1} d\epsilon = \frac{1}{2} \left(\frac{kT}{h\nu} \right)^3 \int_0^{\infty} \frac{x^2}{e^x - 1} dx.$$

This integral is evaluated before, approximately 2.404. Again, the particle number depends on the temperature, which cannot be true. There is only one temperature, the condensation temperature, at which the particle number is true:

$$N = 1.202 \left(\frac{kT_c}{h\nu} \right)^3 \implies T_c = \frac{h\nu}{k} \left(\frac{N}{1.202} \right)^{1/3}.$$

