

# Thermodynamics and Statistical Physics

**B. Zheng**<sup>1</sup>

Zhejiang Institute of Modern Physics, Zhejiang University,  
Hangzhou 310027, P.R. China

---

<sup>1</sup>e-mail: zheng@zimp.zju.edu.cn; tel: 0086-571-87952753; Fax: 0086-571-87952754

# Chapter 1

## Thermodynamics

# Chapter 2

## Classical statistical mechanics

From thermodynamics one does not understand everything. E.g., what is pressure, and especially what are temperature and entropy? Kinetic theory of molecules offers certain understanding, but not sufficiently general. Thermodynamics does not look fundamental and systematic, and all the laws seem isolated each other.

### 2.1 Postulates

Statistical physics

- equilibrium
- non-equilibrium

Statistical mechanics here is mainly concerned with the equilibrium state.

- Calculate macroscopic parameters from microstructures and interactions

- Derive thermodynamics  
e.g., define and calculate  
Temperature  
Internal energy  
Entropy  
Prove that the laws in thermodynamics
- Statistical mechanics goes beyond thermodynamics

The system: a classical system isolated in the sense that the energy is conserved or nearly conserved, is composed of a large number  $N$  of elements, typically

$$N \simeq 10^{23} \quad \text{molecules}$$

Thermodynamic limit

$$N \rightarrow \infty, \quad V \rightarrow \infty$$

but

$$\frac{V}{N} \rightarrow v \quad (\text{a constant})$$

Here  $V$  is measured in a microscopic unit, e.g.,

$$V \simeq 10^{23} \quad \text{molecular volumes}$$

The thermodynamic properties of the system may be tackled from three levels.

- Fundamental level

Solve the microscopic equations of motion, such as Newton, Heisenberg Eqs..

Difficulties

- \* too many degrees of freedom

- \* microscopic initial conditions and boundary conditions
- \* irregular disturbance from environments

Here the time  $t$  is microscopic.

- Quasi-fundamental level

Do not trace the motion of each molecule, and consider only the probability distribution  $\rho(p_i, q_i, t)$  with

$$\rho(p_i, q_i, t) \prod_i d^3 p_i d^3 q_i$$

being the number of molecules at time  $t$ , and lying within a volume  $\prod_i d^3 p_i d^3 q_i$  of the coordinate  $q_i$  and momentum  $p_i$ .

Equations of motion:

- \* Liouville's eq.
- \* Boltzmann eq.

These eqs. can be solved only in some simple cases, such as dilute gases.

The time  $t$  is mesoscopic.

- Statistical mechanics

Do not solve any eqs. of motion, but assume a form of  $\rho(p_i, q_i, t)$  in the equilibrium state, i.e.  $\rho(p_i, q_i, \infty) \equiv \rho(p, q)$

- \* it can be tested by experiments
- \* it can be derived from eqs. of motion in some special cases.

More strictly, this is the so-called ensemble theory.

The time  $t$  is macroscopic.

$\Gamma$  space: the phase space spanned by  $(p, q)$ , each point in  $\Gamma$  space represents a microscopic state of the system.

**An ensemble:** A collection of systems, identical in composition and macroscopic conditions, but existing in different microscopic states, without interactions each other.

A system can be represented by a point in  $\Gamma$  space, then

$$\rho(p, q) dp dq \equiv \rho(p_i, q_i) \prod_i dp_i dq_i$$

is the number of systems at the volume element  $dpdq$ . Physically, it is reasonable to do this just because a macroscopic state may correspond to many microscopic states.

The ensemble is introduced to replace the dynamic evolution of the microscopic states.

**Postulate of equal a priori probability**

When an isolated system is in thermodynamic equilibrium, its state is equally like to be any state satisfying the macroscopic conditions, i.e.

$$\rho(p, q) = \begin{cases} \text{const.} & \text{if } E < H(p, q) < E + \Delta \\ 0 & \text{otherwise} \end{cases}$$

here  $\Delta \ll E$   $H$  is the Hamiltonian and the ensemble described by this distribution is the so-called microcanonical ensemble.

- why is  $\Delta$  introduced?  
Theoretically, it is convenient  
Practically, isolation is not strict
- why is  $\rho(p, q)$  a const for possible states?

One may understand from ergodicity, which may be achieved by

- \* the measure of non-ergodic states is negligibly small
- \* disturbance
- \* boundary conditions
- \* interactions

.....

If there is ergodicity, it is natural that the practical path of a system should be a simple loop in  $\Gamma$  space.

The ensemble average of a measurable property  $f(p, q)$  is defined as

$$\langle f \rangle \equiv \frac{\int dp dq f(p, q) \rho(p, q)}{\int dp dq \rho(p, q)}$$

In the thermodynamic limit, it is usually assumed

$$\frac{\langle f^2 \rangle - \langle f \rangle^2}{\langle f \rangle^2} \ll 1 \quad (*)$$

We may define the most probable value of  $f(p, q)$  as the value with the maximum probability  $P(f_0)$ . Here we should note that  $P(f)$  is different from  $\rho(p, q)$ , as shown in Fig. 2.1.

In the thermodynamic limit, the ensemble average and the most probable value should be nearly the same.

Otherwise, statistical mechanics should be questioned.

Question: in what case Eq. (\*) is not valid?

Answer: strongly correlated systems.

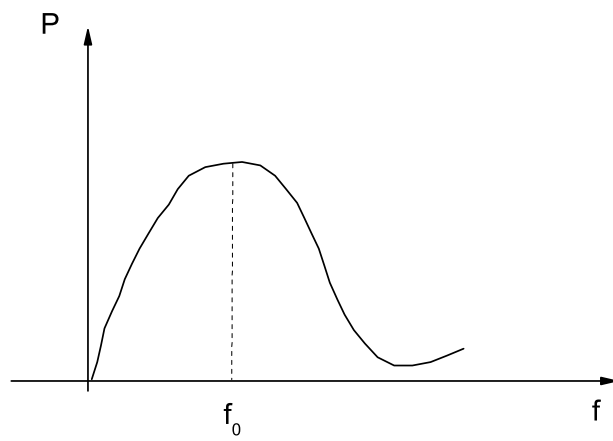


Figure 2.1:



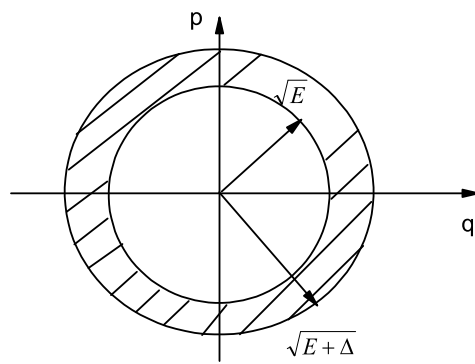


Figure 2.2:

## 2.2 Microcanonical ensemble

$H$  is naturally defined as the internal energy. So, it is important how to define the entropy and temperature.

Let us denote the volume in  $\Gamma$  space of the microscopic ensemble

$$\Gamma(E) \equiv \int_{E < H(p,q) < E+\Delta} dp dq$$

$\Gamma(E)$  is understood to be dependent of  $N, V$  and also  $\Delta$ .

For example, it is shown in Fig.2.2 for  $H(p, q) = p^2 + q^2$ .

The entropy is defined by

$$S(E, V) \equiv k \log \Gamma(E)$$

where  $k$  is a universal constant eventually shown to be Boltzmann's constant.

To justify the definition of  $S$ , we should show

(a)  $S$  is extensive

(b)  $S$  satisfies the properties of the entropy as required by the second law of thermodynamics

**Proof:** (a)

Let the system be divided into two subsystems:

1:  $N_1, V_1, H_1$

2:  $N_2, V_2, H_2$

with

$$N = N_1 + N_2$$

$$V = V_1 + V_2$$

and

$$H(p, q) = H_1(p_1, q_1) + H_2(p_2, q_2)$$

Here it is assumed that the interaction between two subsystems is negligible.  $(p_1, q_1)$  are the coordinates of particles in the system 1, and  $(p_2, q_2)$  are the coordinates of particles in the system 2.

E.g., if the intermolecular potential is finite-range, and the surface-to-volume ratio of each subsystem is negligibly small.

If we define

$$\begin{aligned} S_1(E_1, V_1) &= k \log \Gamma_1(E_1, \Delta) \\ S_2(E_2, V_2) &= k \log \Gamma_2(E_2, \Delta) \end{aligned}$$

and

$$S(E, V) = k \log \Gamma(E, 2\Delta)$$

we should show in the thermodynamic limit

$$S(E, V) = S_1(E_1, V_1) + S_2(E_2, V_2)$$

If  $E = E_1 + E_2$  is a decomposition of the energy, the volume in  $\Gamma$  space of the whole system is

$$\begin{aligned} &\int_{E_1+E_2 < H_1+H_2 < E_1+E_2+2\Delta} dp_1 dp_2 dq_1 dq_2 \\ &= \int_{E_1 < H_1 < E_1+\Delta} dp_1 dq_1 \int_{E_2 < H_2 < E_2+\Delta} dp_2 dq_2 \\ &= \Gamma_1(E_1)\Gamma_2(E_2) \end{aligned}$$

This means that the extensive property of  $S$  has been proven if the two subsystems are added up without interactions.

The key point is that the possible decomposition of the energy  $E = E_1 + E_2$  is not unique when the system is divided into two subsystems.

Let us divide  $E$  into  $E/\Delta$  intervals with  $E_i = i\Delta, i = 1, \dots, E/\Delta$ , then

$$\Gamma(E) = \sum_{i=1}^{E/\Delta} \Gamma_1(E_i)\Gamma_2(E - E_i)$$

We will show or have to show that only one term in the sum is dominating.

Reading materials:

Let the largest term be  $\Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2)$  with

$$\bar{E}_1 + \bar{E}_2 = E$$

then

$$\Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2) \leq \Gamma(E) \leq \frac{E}{\Delta} \Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2)$$

or

$$S_1(\bar{E}_1, V_1) + S_2(\bar{E}_2, V_2) \leq S(E, V) \leq S_1(\bar{E}_1, V_1) + S_2(\bar{E}_2, V_2) + k \log \frac{E}{\Delta}$$

In the thermodynamic limit, we expect

$$\begin{aligned} \log \Gamma_1 &\propto N_1 \\ \log \Gamma_2 &\propto N_2 \quad (*) \\ E &\propto N_1 + N_2 \end{aligned}$$

then

$$S(E, V) = S_1(\bar{E}_1) + S_2(\bar{E}_2) + O(\log(N))$$

In summary, the extensive property of  $S$  is based on that a decomposition  $E = \bar{E}_1 + \bar{E}_2$  of the energy is dominating when the system is divided into two subsystems,

with fixed  $N_1, N_2$  and  $V_1, V_2$ . Such a dominating decomposition is expected from Eqs. (\*).

In thermodynamic equilibrium, the system is homogeneous, therefore

$$N \propto V$$

In other words,  $E \propto N_1 + N_2 = N$  is simply the extensive property of  $E$ . However,  $\log \Gamma_1 \propto N_1 \propto V_1$  in (\*) indicates already that  $S_1$  is extensive.

Therefore, derivation of  $E = \bar{E}_1 + \bar{E}_2$ , Sounds not very convincing. It tells only that if  $S_1$  and  $S_2$  are extensive then  $S$  is also extensive.

End reading materials

Understanding:

Probably, alternatively, we may think  $\bar{E}_1$  and  $\bar{E}_2$  are the averaged energy of the two subsystems. In the thermodynamic limit, the fluctuation  $\delta$  of  $E_1$  around  $\bar{E}_1$  should be much smaller than  $\bar{E}_1$ . Let us take  $\delta < \Delta$ , then only  $\Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2)$  is dominating

$$\Gamma(E) = \sum_{i=1}^{E/\Delta} \Gamma_1(E_i)\Gamma_2(E - E_i)$$

Therefore

$$S(E, V) = S_1(\bar{E}_1, V_1) + S_2(\bar{E}_2, V_2)$$

with  $E = \bar{E}_1 + \bar{E}_2$ .

This also explains why we can add two systems up in equilibrium

Question: why is it only "an understanding"?

**Answer:**  $\Gamma_1(E_i)\Gamma_2(E - E_i)$  is just  $P(E_1 = E_i)$  – the probability  $E_1$  takes the value  $E_i$ . When we assume  $\delta < \Delta$ , it indicates already that all other terms are negligible.

Reading materials:

Why

$$\Gamma(E) = \sum_{i=1}^{E/\Delta} \Gamma_1(E_i)\Gamma_2(E - E_i) \quad ?$$

e.g.  $N = 2$

$$H = \frac{1}{2m} \sum_{i=1}^2 p_i^2 + \frac{1}{2} \omega^2 \sum_{i=1}^2 q_i^2$$

$$H_1 = \frac{1}{2m} p_1^2 + \frac{1}{2} \omega^2 q_1^2$$

$$H_2 = \frac{1}{2m} p_2^2 + \frac{1}{2} \omega^2 q_2^2$$

$$\Gamma(E) = \int_{E < H < E + 2\Delta} dp_1 dq_1 dp_2 dq_2$$

$$\Gamma_1(E_1) = \int_{E_1 < H_1 < E_1 + \Delta} dp_1 dq_1$$

$$\Gamma_2(E - E_1) = \int_{E - E_1 < H_2 < E - E_2 + \Delta} dp_2 dq_2$$

**If  $E_1$  is fixed,**

$$\Gamma(E) = \Gamma_1(E_1)\Gamma_2(E - E_1)$$

**If not**

$$\Gamma(E) = \sum_{i=1}^{E/\Delta} \Gamma_1(E_i)\Gamma_2(E - E_i)$$

since  $E_1$  takes value from 0 to  $E$ , corresponding to different states.

End reading materials

Actually, this implies that  $E = \bar{E}_1 + \bar{E}_2$  maximizes the function  $\Gamma_1(E_1)\Gamma_2(E_2)$  under the restriction  $\delta E = \delta E_1 + \delta E_2 = 0$ , i.e.

$$\delta(\Gamma_1(E_1)\Gamma_2(E_2)) = 0$$

**with**  $\delta E_1 + \delta E_2 = 0$

This leads to

$$\left. \frac{\partial}{\partial E_1} \log \Gamma_1(E_1) \right|_{E_1=\bar{E}_1} = \left. \frac{\partial}{\partial E_2} \log \Gamma_2(E_2) \right|_{E_2=\bar{E}_2}$$

or

$$\left. \frac{\partial S_1(E_1)}{\partial E_1} \right|_{E=\bar{E}_1} = \left. \frac{\partial S_2(E_2)}{\partial E_2} \right|_{E_2=\bar{E}_2}$$

We define the temperature of any system by

$$\frac{\partial S(E, V)}{\partial E} \equiv \frac{1}{T}$$

Then

$$\frac{\partial S_1}{\partial E_1} = \frac{\partial S}{\partial E_2}$$

is simply the zeroth law

$$T_1 = T_2$$

$T$  defined in this way is an intensive parameter, and  $\partial S/\partial E = 1/T$  is also one of the Maxwell relations in thermodynamics.

If  $S$  is correctly defined,  $T$  should be also correct.

## 2.3 Thermodynamics

Let us define

$$\begin{aligned}\sum(E) &= \int_{H(p,q) < E} dp dq \\ \omega(E) &= \frac{\partial \sum(E)}{\partial E}\end{aligned}$$

Then

$$\begin{aligned}\Gamma(E) &= \omega(E)\Delta \\ \Gamma(E) &= \sum(E + \Delta) - \sum(E)\end{aligned}$$

It can be proved that up to an additive const of the order  $O(\log N)$ , the following definitions are equivalent

$$\begin{aligned}S &= k \log \Gamma(E) \\ S &= k \log \omega(E) \\ S &= k \log \sum(E)\end{aligned}$$

why?— question to think about

Keeping in mind that the energy does not fluctuate so much, it is obvious that  $\sum(E) = \sum(\bar{E}_1) \sum(\bar{E}_2)$ .

(b) With the definition

$$S(E, V) = k \log \sum(E)$$

it is easy to show that  $S$  never decreases, i.e., the second law for a thermally isolated system in thermodynamics.

**Proof:** For our system considered, parameters are  $N, E, V$ . By definition of an isolated system,  $N$  and  $E$



can not change,  $V$  can not decrease. Therefore, the second law here is simply stated as that  $S$  is a non-decreasing function of  $V$ . This is obvious, for  $\Sigma(E)$  is a non-decreasing function of  $V$  by its definition.

$S$  is really entropy.

Assuming that the system is changed slowly by coupling the system to external environments. Then it is a quasi-static process

$$\begin{aligned} dS(E, V) &= \left( \frac{\partial S}{\partial E} \right)_V dE + \left( \frac{\partial S}{\partial V} \right)_E dV \\ &= \frac{1}{T} dE + \left( \frac{\partial S}{\partial V} \right)_E dV \end{aligned}$$

Define the pressure of the system to be

$$P \equiv T \left( \frac{\partial S}{\partial V} \right)_E$$

then

$$dS = \frac{1}{T} (dE + PdV)$$

or

$$dE = TdS - PdV$$

This looks like the first law.

Question: Is it reasonable to define

$$P \equiv T \left( \frac{\partial S}{\partial V} \right)_E$$

Hint:

- $P$  is intensive

- Exercise: prove

$$P = - \left( \frac{\partial E}{\partial V} \right)_S$$

If we accept the definition of  $P$ , then  $S$  should be entropy by the first law.

In other words, the first law should be also an assumption in statistical mechanics.

## 2.4 Equipartition Theorem

Reading materials.

$$H = H(p, q) = H(p_j, q_j), \quad j = 1, \dots, 3N$$

e.g.

$$H = \frac{1}{2m} \sum p_i^2$$

or

$$H = \frac{1}{2m} \sum p_i^2 + \frac{1}{2} \omega^2 \sum q_i^2$$

Let  $x_i$  be either  $p_i$  or  $q_i$

$$\begin{aligned} \left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle &= \frac{1}{\Gamma(E)} \int_{E < H < E + \Delta} dp dq x_i \frac{\partial H}{\partial x_j} \\ &= \frac{\Delta}{\Gamma(E)} \frac{1}{\Delta} \left( \int_{H < E + \Delta} dp dq - \int_{H < E} dp dq \right) x_i \frac{\partial H}{\partial x_j} \\ &= \frac{\Delta}{\Gamma(E)} \frac{\partial}{\partial E} \int_{H < E} dp dq x_i \frac{\partial H}{\partial x_j} \end{aligned}$$

$$\int_{H<E} dp dq x_i \frac{\partial H}{\partial x_j} = \int_{H<E} dp dq x_i \frac{\partial}{\partial x_j} (H - E)$$

$$= \int_{H<E} dp dq \frac{\partial}{\partial x_j} [x_i (H - E)] - \delta_{ij} \int_{H<E} dp dq (H - E)$$

the first term =  $\int_{H=E} dp dq x_i (H - E) = 0$

$$\therefore \left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle = \frac{\delta_{ij}}{\omega(E)} \frac{\partial}{\partial E} \int_{H<E} dp dq (E - H)$$

$$= \frac{\delta_{ij}}{\omega(E)} \left( \int_{H<E} dp dq + \frac{1}{\Delta} \int_{E<H<E+\Delta} dp dq (E - H) \right)$$

(the second term is negligible)

$$= \frac{\delta_{ij}}{\omega(E)} \sum(E) = \delta_{ij} \frac{1}{\frac{\partial}{\partial E} \log \sum(E)} = \delta_{ij} \frac{k}{\frac{\partial S}{\partial E}}$$

$$= \delta_{ij} kT$$

**If**  $i = j$

$$\left\langle x_i \frac{\partial H}{\partial x_i} \right\rangle = kT$$

**If**

$$H = \sum_i A_i P_i^2 + \sum_i B_i Q_i^2$$

$P_i, Q_i$  are canonical conjugate variables

**Then**

$$\sum_i \left( P_i \frac{\partial H}{\partial P_i} + Q_i \frac{\partial H}{\partial Q_i} \right) = 2H$$

**If**  $f$  of the const  $A_i$  and  $B_i$  are non-zero,

$$\langle H \rangle = \frac{1}{2} f kT$$

exercise

Prove the equipartition theorem for

$$H = \frac{1}{2m}p^2 + \frac{1}{2}\omega^2q^2$$

by explicitly calculating the ensemble average.

## 2.5 Classical ideal gas

$$\begin{aligned} H &= \frac{1}{2m} \sum_{i=1}^N p_i^2 \\ \sum(E) &= \int_{H<E} d^3p_1 \cdots d^3p_N d^3q_1 \cdots d^3q_N \\ &= V^N \int_{H<E} d^3p_1 \cdots d^3p_N \end{aligned}$$

Let

$$\begin{aligned} R &= \sqrt{2mE} \\ \sum(E) &= V^N \Omega_{3N}(R) \end{aligned}$$

$\Omega_n(R)$  is the volume of an n-sphere of radius  $R$

$$\begin{aligned} \Omega_n(R) &= C_n R^n \\ C_n &= \frac{2\pi^{n/2}}{\Gamma(n/2 + 1)} \end{aligned}$$

$\Gamma(Z)$  is the gamma function

$$\log C_n \xrightarrow{n \rightarrow \infty} \frac{n}{2} \log \pi - \frac{n}{2} \log \frac{n}{2} + \frac{n}{2}$$

hence

$$\begin{aligned}\sum(E) &= C_{3N} \left[ V(2mE)^{3/2} \right]^N \\ S(E, V) &= k \log \sum(E) \\ &\simeq Nk \log \left[ V \left( \frac{4\pi m E}{3N} \right)^{3/2} \right] + \frac{3}{2} Nk\end{aligned}$$

$$\begin{aligned}U(S, V) \equiv E &= \frac{3N}{4\pi m V^{2/3}} \exp \left( \frac{2S}{3Nk} - 1 \right) \\ T &= \left( \frac{\partial U}{\partial S} \right)_V = \frac{2U}{3Nk}, \quad U = \frac{3}{2} NkT \\ C_V &= \frac{3}{2} Nk \\ P &= - \left( \frac{\partial U}{\partial V} \right)_S = \frac{NkT}{V}\end{aligned}$$

## 2.6 Maxwell-Boltzmann distribution

$N$  identical molecules, volume  $V$ ,

$$H(p, q) = E$$

The system may be described by a microcanonical ensemble. For quasi-independent systems:

$$H = \sum_i H_i(p_i, q_i)$$

where  $H_i(p_i, q_i)$  represents the Hamiltonian of each molecule. There could be a number of atoms in a molecule, but

the interactions between two molecules are negligible. E.g., the simplest case is

$$H = \frac{1}{2m} \sum_i p_i^2$$

$\mu$  space: the phase space spanned by the single-molecule coordinates  $(p, q)$ .

A microscopic state of a single molecule can be represented by a point in the  $\mu$  space. A microscopic state of the system is described by a set of the points.

Since the energy of a molecule is bounded by  $E$ , the points are confined to a finite region of  $\mu$  space. We divide the region into  $K$  elements of volume  $\omega = d^3p d^3q$ , and denote the number of the molecules in a element by  $n_l$ , then

$$\sum_{l=1}^K n_l = N \quad \sum_{l=1}^K \epsilon_l n_l = E.$$

Here the molecules are assumed to be quasi-independent,  $\epsilon_l$  is the energy level of a single particle.

It is important to note that a microscopic state of the system may be described by a set of  $\{n_l\}$ , but a set of  $\{n_l\}$  corresponds to not only one microscopic state, e.g., interchange of two molecules leads to new states. That is, a given set of  $\{n_l\}$  corresponds to a volume in  $\Gamma$  space, which is called the volume occupied by  $\{n_l\}$ .

To describe a macroscopic state, we need to average over the microcanonical ensemble, i.e., all possible microscopic states. In other words, we should obtain  $\{\langle n_l \rangle\}$ , then calculate all macroscopic quantities.

However, it is difficult to perform this average.

We assume that the equilibrium state is described by the *most probable distribution*  $\{n_l\}$ , which occupies the largest volume in  $\Gamma$  space.

Why?

- This is some what similar to the case when we calculate  $\Gamma(E) \simeq \Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2)$  by dividing the system into two subsystems.
- For a fixed  $l$ , if the relative fluctuation

$$(\langle n_l^2 \rangle - \langle n_l \rangle^2) / \langle n_l \rangle^2$$

is still not small enough, we increase the total number  $N$  to reduce it. Finally,  $\langle n_l \rangle$  should be equal to the most probable  $n_l$ .

The procedure:

- (a) Calculate the volume of  $\{n_l\}$
- (b) Maximize it to obtain the most probable  $\{n_l\}$ .

The volume of  $\{n_l\}$

$$\Omega(\{n_l\}) \propto \frac{N!}{\prod_{l=1}^K n_l!} \prod_{k=1}^K g_k^{n_k}$$

where  $g_k$  is introduced for convenience and will finally be put to 1.

Understanding:

There are  $N!$  ways of distributing  $N$  distinguishable molecules to  $N$  positions. However,  $N$  positions form  $K$  groups with the distribution  $\{n_l\}$ . Inside a group, there are  $n_l!$  ways of distributing  $n_l$  molecules.

For a large  $n_l$ ,

$$\log n_l! = n_l(\log n_l - 1) \quad (\text{i.e., } n_l! \simeq n_l^{n_l})$$

$$\begin{aligned} \log \Omega(\{n_l\}) &= N(\log N - 1) - \sum_{l=1}^K n_l(\log n_l - 1) \\ &\quad + \sum_{l=1}^K n_l \log g_l + \text{const} \end{aligned}$$

Now we vary  $\{n_l\}$  under the condition of  $\sum_{l=1}^K n_l = N$  and  $\sum_{l=1}^K \epsilon_l n_l = E$ , to find the most probable  $\{n_l\}$ . We introduce the Lagrange multipliers  $\alpha$  and  $\beta$ , and calculate

$$\delta [\log \Omega(\{n_l\})] - \delta \left( \alpha \sum_{l=1}^K n_l + \beta \sum_{l=1}^K \epsilon_l n_l \right) = 0$$

Now we consider all  $n_l$  are independent each other.

$$\sum_{l=1}^K [-(\log n_l) + \log g_l - \alpha - \beta \epsilon_l] \delta n_l = 0$$

$$\begin{aligned} \therefore \quad \log \bar{n}_l &= -\alpha - \beta \epsilon_l \\ \bar{n}_l &= e^{-\alpha - \beta \epsilon_l} \end{aligned}$$

Finally,  $\alpha$  and  $\beta$  can be determined by the conservation of the total particles and the total energy.

To prove that  $\{n_l\}$  maximizes  $\Omega(\{n_l\})$ , we can simply calculate the second variation

$$- \sum_{l=1}^K \frac{1}{n_l} (\delta n_l)^2 < 0$$



Note that

\*  $\bar{n}_l$  is only the function of  $\epsilon_l$ , does not generally depends on  $(p, q)$ . This is important in the equilibrium state.

\* The molecules tends to gather in the lower energy states.

## 2.7 Boltzmann statistical theory

$$N = \sum_l \bar{n}_l = \sum_l e^{-\alpha - \beta \epsilon_l}$$
$$E = \sum_l \epsilon_l \bar{n}_l = \sum_l \epsilon_l e^{-\alpha - \beta \epsilon_l}$$

Let us assume

$$\epsilon_l = \epsilon(p, q, y)$$

$y = \{y_k\}$  represent macroscopic external parameters.

Define the partition function of a single particle

$$Z(\beta, y) = \sum_l e^{-\beta \epsilon_l} = \int_{\epsilon \leq E} dpdq e^{-\beta \epsilon(p, q, y)}$$

then

$$N = e^{-\alpha} Z(\beta, y) \quad \text{or} \quad \alpha = \log \frac{Z(\beta, y)}{N}$$
$$E = -N \frac{\partial \log Z(\beta, y)}{\partial \beta}$$

Suppose the system is changed very slowly.

$$dE = \sum_l \bar{n}_l d\epsilon_l + \sum_l \epsilon_l d\bar{n}_l$$

The first term represents the interaction with the external environment

$$\begin{aligned} \sum_l \bar{n}_l d\epsilon_l &= \sum_{l,k} \bar{n}_l \frac{\partial \epsilon_l}{\partial y_k} dy_k \\ &= \sum_k \left( \sum_l \bar{n}_l \frac{\partial \epsilon_l}{\partial y_k} \right) dy_k \end{aligned}$$

**Question:** why not differentiate over  $p$  and  $q$ ?

**Answer:**  $p$  and  $q$  are integrating variables.

Here

$$Y_k \equiv \sum_l \bar{n}_l \frac{\partial \epsilon_l}{\partial y_k} = -\frac{N}{\beta} \frac{\partial \log Z(\beta, y)}{\partial y_k}$$

are the forces acting on the system from the environment.

**Proof:** since

$$e^{-\alpha} = \frac{N}{Z(\beta, y)}$$

therefore

$$\begin{aligned} &-\frac{N}{\beta} \frac{\partial \log Z(\beta, y)}{\partial y_k} \\ &= -\frac{N}{\beta} \frac{1}{Z(\beta, y)} \frac{\partial Z(\beta, y)}{\partial y_k} \\ &= -\frac{e^{-\alpha}}{\beta} \frac{\partial}{\partial y_k} \sum_l e^{-\beta \epsilon_l} \\ &= \sum_l e^{-\alpha - \beta \epsilon_l} \frac{\partial \epsilon_l}{\partial y_k} \end{aligned}$$

**For example,  $y_k = -V$ , then**

$$P = +Y = \frac{N}{\beta} \frac{\partial \log Z(\beta, V)}{\partial V}$$

**If**

$$\begin{aligned} d\epsilon_l &= 0 \\ dQ &= \sum_l \epsilon_l d\bar{n}_l \end{aligned}$$

**Therefore, in general**

$$\begin{aligned} dQ &= \sum_l \epsilon_l d\bar{n}_l = dE - \sum_k Y_k dy_k \\ &= -Nd \left( \frac{\partial \log Z(\beta, y)}{\partial \beta} \right) + \frac{N}{\beta} \sum_k \frac{\partial \log Z(\beta, y)}{\partial y_k} dy_k \end{aligned}$$

**Since**

$$d \log Z(\beta, y) = \frac{\partial \log Z(\beta, y)}{\partial \beta} d\beta + \sum_k \frac{\partial \log Z(\beta, y)}{\partial y_k} dy_k$$

**hence**

$$dQ = \frac{N}{\beta} d \left( \log Z(\beta, y) - \beta \frac{\partial \log Z(\beta, y)}{\partial \beta} \right)$$

**Define**

$$\beta = \frac{1}{kT}$$

$$\begin{aligned} dS &= Nk d \left( \log Z(\beta, y) - \beta \frac{\partial \log Z(\beta, y)}{\partial \beta} \right) \\ \text{i.e., } S &= Nk \left( \log Z(\beta, y) - \beta \frac{\partial \log Z(\beta, y)}{\partial \beta} \right) + \text{const} \end{aligned}$$

Then,  $dQ = TdS$ , i.e., assuming the first law, we may prove the second law.

**Exercise:** Perform the calculations for an ideal gas with

$$H = \frac{1}{2m} \sum_i p_i^2$$

**Hint:** Since  $E \propto N$ , we neglect the bound of  $E$  in calculating

$$Z(\beta, y) = \sum_l e^{-\beta \epsilon_l}$$

since  $\epsilon_l \ll E$ .

**Question:** What is the relation between the Boltzmann theory and the microcanonical theory?

**Answer:**

$$\begin{aligned} \log \Omega(\{\bar{n}_l\}) &= N \log N - \sum_l \bar{n}_l \log \bar{n}_l \\ &= N \log N + \sum_l (\alpha + \beta \epsilon_l) e^{-\alpha - \beta \epsilon_l} \\ &= N \log N + \alpha \sum_l e^{-\alpha - \beta \epsilon_l} + \beta \sum_l \epsilon_l e^{-\alpha - \beta \epsilon_l} \\ &= N \log N + N \log \frac{Z(\beta, y)}{N} - N \beta \frac{\partial \log Z(\beta, y)}{\partial \beta} \\ \therefore S &= Nk \left( \log Z(\beta, y) - \beta \frac{\partial \log Z(\beta, y)}{\partial \beta} \right) \end{aligned}$$

**Fix  $y$ ,**

$$dS = Nk \frac{\partial \log Z(\beta, y)}{\partial \beta} d\beta - Nk \frac{\partial \log Z(\beta, y)}{\partial \beta} d\beta + k\beta dE$$

$$\therefore \frac{1}{T} = \left( \frac{\partial S}{\partial E} \right)_y = k\beta$$

## 2.8 Summary

### Ergodicity

→ the probability distribution of a microcanonical ensemble

A special case: the most probable distribution of quasi-independent particles.

→  $T, S, P$  and thermodynamics

The drawback is that the calculation of observables is clumsy because of the restriction of the energy.

**Exercise**:

Assuming

$$P \equiv T \left( \frac{\partial S}{\partial V} \right)_E, \quad \left( \frac{\partial S}{\partial E} \right)_V = \frac{1}{T},$$

prove

$$P = - \left( \frac{\partial E}{\partial V} \right)_S.$$

Prove the equipartition theorem for

$$H = \frac{1}{2m}p^2 + \frac{1}{2}\omega^2q^2$$

by explicitly calculating the ensemble average.

Perform the calculations of the Boltzmann statistical theory for an ideal gas with

$$H = \frac{1}{2m} \sum_i p_i^2$$

Exercises 6.1 and 6.3 in the textbook.