

Notes

Statistical Mechanics

Yi Huang¹

University of Minnesota

E-mail: yihphysics@gmail.com

¹ *[My Personal Website](#)*

Contents

I	Caprice	1
1	Fall 2018	1
1.1	Proof of (1.4.12)	1
1.2	Macroscopic systems are in general in mixed states	1
1.3	Problem 1.14	1
1.4	Counter Intuition Fact about High-Dimensional Ball	1
1.5	Derivation of (1.3.6)	3
1.6	Proof of (2.3.5)	4
1.7	How does statistical mechanics relate to probability theory?	4
1.8	Derivation of (2.6.3)	4
1.9	Derivation of (2.6.12)	5
1.10	An interesting derivative	5
1.11	Statistical Ensemble	5
1.12	Density matrix in Schrödinger picture and Heisenberg picture	5
1.13	Two equivalent statement of equilibrium statistical mechanics hypothesis	5
1.14	A Thought	6
1.15	TdS equation	6
1.16	Legendre transformation	6
1.17	Property of δ function	6
1.18	The partition function of ideal quantum gas	7
1.19	The average occupation number of ideal quantum gas	9
1.20	Maxwell-Boltzmann statistics	10
1.21	Riemann Zeta function	11

Caprice

PART

I

1 Fall 2018

This section is based on my course of statistical mechanics in Fall 2018 taught by Professor Kapusta. The textbook we use is Schwabl's *Statistical Mechanics*. I also refer to Bellac's *Equilibrium and Non-Equilibrium Statistical Thermodynamics*. | Section 1. Fall 2018

1.1 Proof of (1.4.12)

Proof. We use the fact that trace of a matrix is invariant under change of basis.

$$\begin{aligned}\langle A \rangle_t &= \text{Tr}(\rho(t)A) \\ &= \text{Tr}(U\rho(t_0)U^\dagger A) \\ &= \text{Tr}(U^\dagger(U\rho(t_0)U^\dagger A)U) \\ &= \text{Tr}(\rho(t_0)U^\dagger AU) \\ &= \text{Tr}(\rho(t_0)A_H(t))\end{aligned}$$

□

1.2 Macroscopic systems are in general in mixed states

Consider a system consisting of two subsystems 1 and 2, and in a pure state

$$|\psi\rangle = \sum_{n,m} c_{nm} |1n\rangle |2m\rangle \quad (1.1)$$

1.3 Problem 1.14

Let $Y = \ln X$ be a random variable obeys a Gaussian distribution, then $X = X(Y) = e^Y$ is a function of Y , with the probability density

$$\begin{aligned}\omega_X(x) &= \langle \delta(X(Y) - x) \rangle \\ &= \int dy \omega_Y(y) \delta(e^y - x) \\ &= \int dy \omega_Y(y) \frac{\delta(y - \ln x)}{|x|} \\ &= \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln(x/x_0))^2}{2\sigma^2}}, \quad 0 < x < \infty\end{aligned}$$

1.4 Counter Intuition Fact about High-Dimensional Ball

Nearly the whole volume of a high-dimensional ball lies at its surface.

The volume of a N dimensional ball

$$V_N(r) = \frac{\pi^{N/2}}{\Gamma(\frac{N}{2} + 1)} r^N, \quad (1.2)$$

The surface area of a N dimensional ball (area of S_N surface)

$$S_N(r) = \frac{2\pi^{N/2}}{\Gamma(\frac{N}{2})} r^{N-1}, \quad (1.3)$$

Suppose the thickness of this hypersphere is Δr , set the shell volume equal to the total ball volume to find the required Δr

$$\begin{aligned} S_N(r)\Delta r &= V_N(r), \\ \Delta r &= \frac{V_N(r)}{S_N(r)} = \frac{r}{N}. \end{aligned}$$

For finite radius r , as the dimension gets higher and higher, say $N \rightarrow \infty$, then the ratio between the thickness and the radius goes to zero,

$$\frac{\Delta r}{r} = \frac{1}{N} \rightarrow 0, \quad (1.4)$$

which means almost all the volume of this hypersphere lies on its surface.

As for how to calculate the volume and the surface area of a N -dimensional ball, we can consider a Gaussian integral

$$G(N) = \int dx_1 dx_2 \dots dx_N \exp\left(-\sum_{i=1}^N x_i^2\right), \quad (1.5)$$

Actually it is easy to get the result

$$G(N) = \left(\int_{-\infty}^{\infty} dx e^{-x^2}\right)^N = \pi^{N/2}, \quad (1.6)$$

However we can also compute this integral in polar coordinate

$$\begin{aligned} G(N) &= \int_0^{\infty} dr \int_{S_N(r)} dS_N(r) e^{-r^2} \\ &= \int_0^{\infty} dr S_N(r) e^{-r^2}, \end{aligned}$$

Although we don't know what is $S_N(r)$, what we do know is

$$V_N(r) = V_N(1)r^N, \quad (1.7)$$

$$S_N(r) = \frac{dV_N(r)}{dr} = NV_N(1)r^{N-1}. \quad (1.8)$$

Substitute into $G(N)$

$$\begin{aligned} G(N) &= NV_N(1) \int_0^{\infty} dr r^{N-1} e^{-r^2} \\ &= \frac{1}{2} NV_N(1) \int_0^{\infty} d(r^2) (r^2)^{N/2-1} e^{-r^2} \\ &= \frac{1}{2} NV_N(1) \int_0^{\infty} dz z^{N/2-1} e^{-z} \\ &= \frac{1}{2} NV_N(1) \Gamma(N/2). \end{aligned}$$

Here we use the definition of Gamma function

$$\Gamma(t) = \int_0^\infty dz z^{t-1} e^{-z}. \quad (1.9)$$

Thus we have

$$\frac{1}{2} N V_N(1) \Gamma(N/2) = \pi^{N/2}, \quad (1.10)$$

which means

$$V_N(r) = \frac{\pi^{N/2}}{\Gamma(N/2 + 1)} r^N, \quad (1.11)$$

and

$$S_N(r) = \frac{2\pi^{N/2}}{\Gamma(N/2)} r^{N-1} \quad (1.12)$$

1.5 Derivation of (1.3.6)

Classically, the probability density in phase space is given by

$$\rho(q, p, t) = \int dq_0 dp_0 W(q_0, p_0) \delta(q - q(t; q_0, p_0)) \delta(p - p(t; q_0, p_0)). \quad (1.13)$$

We want to calculate time derivative of ρ

$$\frac{\partial \rho}{\partial t} = \int dq_0 dp_0 W(q_0, p_0) \frac{\partial}{\partial t} [\delta(q - q(t; q_0, p_0)) \delta(p - p(t; q_0, p_0))]. \quad (1.14)$$

To compute this partial derivative, we use a property of derivative

$$\frac{d}{dx} v(x) = -\frac{d}{dx} v(-x), \quad (1.15)$$

and

$$\frac{d}{d(x-L)} v(x-L) = \frac{d}{dx} v(x-L) = \frac{d}{dL} v(L-x) = -\frac{d}{dL} v(x-L). \quad (1.16)$$

Thus

$$\begin{aligned} f(x, t) &= \int dx_0 \omega(x_0) \delta(x - q_x(t; x_0)) \\ \frac{\partial}{\partial t} f(x, t) &= \int dx_0 \omega(x_0) \frac{\partial(x - q_x(t; x_0))}{\partial t} \frac{\partial \delta(x - q_x(t; x_0))}{\partial(x - q_x(t; x_0))} \\ &= - \int dx_0 \omega(x_0) \dot{q}_x(t; x_0) \frac{\partial \delta(x - q_x(t; x_0))}{\partial(x - q_x(t; x_0))} \\ &= - \int dx_0 \omega(x_0) \dot{q}_x \frac{\partial}{\partial x} \delta(x - q_x(t; x_0)) \\ &= -\dot{x} \frac{\partial}{\partial x} \int dx_0 \omega(x_0) \delta(x - q_x(t; x_0)) \\ &= -\dot{x} \frac{\partial}{\partial x} f(x, t). \end{aligned}$$

Thus

$$\frac{\partial \rho}{\partial t} = - \sum_i \left(\dot{q}_i \frac{\partial}{\partial q_i} + \dot{p}_i \frac{\partial}{\partial p_i} \right) \rho \quad (1.17)$$

1.6 Proof of (2.3.5)

Entropy is defined by

$$S[\rho] = -k \text{Tr}(\rho \ln \rho) = -k \langle \ln \rho \rangle. \quad (1.18)$$

The entropy of microcanonical ensemble S_{MC} is largest element of the entropy of any given density matrix. There is a theorem (2.3.3)

$$\text{Tr}(\rho(\ln \rho_1 - \ln \rho)) < 0, \quad (1.19)$$

if $\rho \neq \rho_1$.

For microcanonical ensemble, $S[\rho_{MC}] = \text{Tr}(\rho_{MC} \ln \rho_{MC}) = \text{Tr}(\rho \ln \rho_{MC})$ is true for any density matrix ρ .

So

$$S[\rho] = -k \text{Tr}(\rho \ln \rho) \leq -k \text{Tr}(\rho \ln \rho_{MC}) = S_{MC}, \quad (1.20)$$

$S[\rho] = S_{MC}$ if and only if $\rho = \rho_{MC}$, which means S_{MC} is the largest element of entropy.

Notice that $-k \text{Tr}(\rho \ln \rho_1)$ doesn't have physical meaning if $\rho \neq \rho_1$ but it is also an upper bound of $S[\rho]$, since

$$S[\rho] = -k \text{Tr}(\rho \ln \rho) < -k \text{Tr}(\rho \ln \rho_1) \quad (1.21)$$

1.7 How does statistical mechanics relate to probability theory?

The random variables in statistical mechanics are $3N$ coordinates q_i and $3N$ canonical momentum p_i , $i = 1, 2, \dots, 3N$. Also functions of these $6N$ random variables are also random variables, for example, the Hamiltonian $H(q, p)$.

Why are the $6N$ degrees of freedom random variables?

Because we cannot solve for a macroscopic system exactly in general when $N \rightarrow \infty$, nor even know the initial conditions. We should expect there is a distribution function to assign probability to every point in the phase space. That's where the density matrix comes from.

1.8 Derivation of (2.6.3)

\tilde{E}_1 is the energy of subsystem 1 when the whole system is in the most probable state. Suppose when the overall system is at equilibrium state, $(\tilde{E}_1 - E_{1n})$ is much small compare to the energy of the heat bath $(E - \tilde{E}_1)$, so we can expand $\ln \Omega_2$

$$\begin{aligned} \ln \Omega_2(E - E_{1n}) &= \ln \Omega_2(E - \tilde{E}_1 + \tilde{E}_1 - E_{1n}) \\ &= \ln \Omega_2(E - \tilde{E}_1) + \left[\frac{\partial}{\partial (E - \tilde{E}_1)} \ln \Omega_2(E - \tilde{E}_1) \right]_{E - \tilde{E}_1} (\tilde{E}_1 - E_{1n}) \\ &= \ln \Omega_2(E - \tilde{E}_1) + \frac{\tilde{E}_1 - E_{1n}}{kT} \end{aligned}$$

Therefore

$$\Omega_2(E - E_{1n}) = \Omega_2(E - \tilde{E}_1) \exp \left(\frac{\tilde{E}_1 - E_{1n}}{kT} \right). \quad (1.22)$$

1.9 Derivation of (2.6.12)

The partition function of canonical ensemble is

$$\begin{aligned}
 Z &= \text{Tr}_1 \exp(-H_1/kT) \\
 &= \int dE_1 \text{Tr}_1 \exp(-H_1/kT) \delta(H_1 - E_1) \\
 &= \int dE_1 \text{Tr}_1 \exp(-E_1/kT) \delta(H_1 - E_1) \\
 &= \int dE_1 \exp(-E_1/kT) \Omega_1(E_1).
 \end{aligned}$$

1.10 An interesting derivative

Suppose $v(x_1 - x_2)$ is a function of relative position, then

$$\left(x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} \right) v(x_1 - x_2) = (x_1 - x_2) \frac{\partial v(x_1 - x_2)}{\partial (x_1 - x_2)} \quad (1.23)$$

1.11 Statistical Ensemble

Remark. The collection of all the microstates which represent a macrostate, weighted by their frequency of occurrence, is referred to as a statistical ensemble. - page 3 in Schwabl's textbook.

In this statement, he assumes that the frequency of occurrence of the time scale of measurement is equal to the probability of the microstates. Which means the number of microstate transitions during a measurement is really large. Suppose the microstate transition is driven by external perturbation, then this large number of transitions means the external perturbation happens at a pretty fast rate.

1.12 Density matrix in Schrödinger picture and Heisenberg picture

The density matrix ρ actually represents the property of states themselves, so in Schrödinger picture, ρ is time-dependent, such that

$$\begin{aligned}
 \rho(t) &= \sum_i p_i |\psi_i(t)\rangle \langle \psi_i(t)| \\
 &= \sum_i p_i U(t, t_0) |\psi_i\rangle \langle \psi_i| U^\dagger(t, t_0) \\
 &= U(t, t_0) \rho(t_0) U^\dagger(t, t_0)
 \end{aligned}$$

However, in Heisenberg picture, ρ should be time-independent.

1.13 Two equivalent statement of equilibrium statistical mechanics hypothesis

1. An equilibrium state requires that

$$\dot{\rho} = \frac{\partial \rho}{\partial t} = -\frac{1}{i\hbar} [\rho, H] = 0, \quad (1.24)$$

2. All microstates in the energy shell are equally probable.

Next we prove they are equivalent, given the fact that ergodic hypothesis is true in classical case.

- $1 \rightarrow 2$: Classically, $\frac{\partial \rho}{\partial t} = 0$ has the physical meaning that for a fixed position in phase space, the probability density doesn't change as time goes by. Since the probability density are like incompressible flow and if you go with the trajectory in phase space, the probability density doesn't change according to Liouville theorem $\dot{\rho} = 0$, the fact that you observe constant density at a fixed position implies before the probability density flows into this position, its magnitude is also the same constant. If ergodic hypothesis is true, the the phase trajectory pass by this position will "fill" the whole energy shell, such that the probability density for any position in this energy shell is equal. Quantum mechanically, If regions within the energy shell had different statistical weights, then the distribution function (density matrix) would depend on other quantities besides $H(q, p)$, and ρ would not commute with H (classically, the Poisson bracket would not vanish).
- $2 \rightarrow 1$: Classically, all states within the energy shell are equally probable means ρ is inversely proportional to the volume of the energy shell, which is specified by the Hamiltonian H , and ρ is a function of H . This implies the Poisson bracket $[\rho, H] = 0$. Quantum mechanically, the density matrix sums over the states within the energy shell, which is also specified by the Hamiltonian H . By the similar reason the commutator $[\rho, H] = 0$.

1.14 A Thought

The classical mechanics are discovered in macroscopic world, where we cannot see all microscopic "objects", so it is kind of emergence phenomena or effective theory. However, when we want to write down the equation for microscopic objects, we borrow the classical ideas, for example the classical Hamiltonian, but why that is applicable? Is that amazing that the emergent classical effective theory can also describe the microscopic objects?

1.15 TdS equation

The following equation of heat is valid only for quasi-static process

$$\delta Q = TdS. \quad (1.25)$$

An infinitesimal reversible process doesn't change the entropy, thus $\delta Q = 0$. This is not true for a reversible process that forms a circle, and in this case, we require $\Delta S = 0$.

1.16 Legendre transformation

Since Legendre transformation only makes sense when the function is convex or concave, does it mean that in mechanics, Lagrangian and Hamiltonian must be convex or concave?

1.17 Property of δ function

Theorem 1.1.

$$\int f(x)\delta(g(x))dx = \sum_i \frac{f(x_i)}{|g'(x_i)|}, \quad (1.26)$$

where x_i are the roots of $g(x) = 0$.

Proof. Change integration variable

$$\begin{aligned} \int_{g=-\infty}^{g=\infty} f(x) \delta(g(x)) \frac{dx}{dg(x)} dg(x) &= \int_{g=-\infty}^{g=\infty} \frac{f(x)}{g'(x)} \delta(g(x)) dg(x) \\ &= \sum_i \int_{g(x_i-\Delta)}^{g(x_i+\Delta)} \frac{f(x)}{g'(x)} \delta(g(x)) dg(x), \end{aligned} \quad (1.27)$$

where in (1.27) $0 < \Delta \ll \min_{i \neq j} \{|x_i - x_j|\}$. If $g'(x_i) > 0$, then $g(x_i + \Delta) > g(x_i - \Delta)$, so we obtain

$$\int_{g(x_i-\Delta)}^{g(x_i+\Delta)} \frac{f(x)}{g'(x)} \delta(g(x)) dg(x) = \frac{f(x_i)}{g'(x_i)}, \quad (1.28)$$

If $g'(x_i) < 0$, then $g(x_i + \Delta) < g(x_i - \Delta)$, so we obtain

$$\int_{g(x_i-\Delta)}^{g(x_i+\Delta)} \frac{f(x)}{g'(x)} \delta(g(x)) dg(x) = -\frac{f(x_i)}{g'(x_i)}. \quad (1.29)$$

From above, we conclude that

$$\int f(x) \delta(g(x)) dx = \sum_i \frac{f(x_i)}{|g'(x_i)|}. \quad (1.30)$$

□

1.18 The partition function of ideal quantum gas

Consider a system of N noninteracting identical particles, the Hamiltonian for the whole system is given by

$$H = \sum_{i=1}^N H_i, \quad (1.31)$$

where H_i is the single-particle Hamiltonian for the i th particle. Suppose the energy spectrum for the single-particle Hamiltonian is ε_l , where l labels the value of the energy and all other quantum numbers necessary to specify the quantum state (see Bellac page 268). That is to say, energy levels that are degenerate but specified by different quantum numbers will be labelled by different values for l . Furthermore, for convenience, we choose the ground state level as the reference energy $\varepsilon_l = 0$. For identical particles, the energy level of the whole system E is specified by the occupation number n_l for each single-particle energy level ε_l , i.e., the number of particles in each energy level.

$$E_r = E\{n_l\} = \sum_{l=0}^{\infty} n_l \varepsilon_l \quad (1.32)$$

where we denote the configuration of particles as $r = \{n_l\}$. The canonical partition function for N identical independent particles subjected to the constraint $N = \sum_l n_l$ is then written as

$$\begin{aligned} Z_N &= \sum_r \exp(-\beta E_r) \\ &= \sum_{\substack{\{n_l\} \\ \sum_l n_l = N}} \exp(-\beta E\{n_l\}) = \sum_{\substack{\{n_l\} \\ \sum_l n_l = N}} \prod_{l=0}^{\infty} \exp(-\beta n_l \varepsilon_l) \end{aligned} \quad (1.33)$$

If for some state ε_l there is no particle occupied at all, then $n_l = 0$ and in the product $\prod_{l=0}^{\infty}$ we are multiplying by 1. In practice, the constraint $N = \sum_l n_l$ is not easily taken into account. Therefore, instead of using the canonical partition function (1.33) it is more convenient to use the grand partition function where the number of particles is not fixed.

$$\begin{aligned} Z_G(T, V, \mu) &= \sum_{N=0}^{\infty} Z_N(V, T) \exp(-\beta\mu N) = \sum_{N=0}^{\infty} z^N Z_N(V, T) \\ &= \sum_{N=0}^{\infty} \sum_{\substack{\{n_l\} \\ \sum_l n_l = N}} \prod_{l=0}^{\infty} [z^{n_l} \exp(-\beta n_l \varepsilon_l)] \end{aligned} \quad (1.34)$$

This expression (1.34) can be simplified by exchanging the order of two summations.

$$\begin{aligned} \sum_{N=0}^{\infty} \sum_{\substack{\{n_l\} \\ \sum_l n_l = N}} &\equiv \sum_{N=0}^{\infty} \sum_{\substack{\{n_l\} \\ \text{no constraint}}} \delta_{N, \sum_l n_l} \\ &\equiv \sum_{\substack{\{n_l\} \\ \text{no constraint}}} \sum_{N=0}^{\infty} \delta_{N, \sum_l n_l} \equiv \sum_{\substack{\{n_l\} \\ \text{no constraint}}} \equiv \sum_{\{n_l\}} \end{aligned} \quad (1.35)$$

For a fixed configuration $\{n_l\}$, the Kronecker delta contributes for only one value of $N = \sum n_l$, such that we can eliminate the summation over N . In (1.35) we abbreviate the summation of all possible configuration $\{n_l\}$ without constraint as $\sum_{\{n_l\}}$. Also we can write it as

$$\sum_{\{n_l\}} \equiv \prod_{l=0}^{\infty} \sum_{\text{All allowed } n_l} \quad (1.36)$$

for bosons

$$\sum_{\{n_l\}} = \sum_{n_0=0}^{\infty} \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \dots \quad (1.37)$$

for fermions

$$\sum_{\{n_l\}} = \sum_{n_0=0}^1 \sum_{n_1=0}^1 \sum_{n_2=0}^1 \dots \quad (1.38)$$

Therefore we can rewrite (1.34)

$$Z_G = \sum_{\{n_l\}} \prod_l (ze^{-\beta\varepsilon_l})^{n_l} \quad (1.39)$$

The grand partition function is also given by the trace

$$\begin{aligned} Z_G &= \text{Tr} \left[e^{-\beta(\hat{H} - \mu\hat{N})} \right] = \sum_r \langle r | e^{-\beta(\hat{H} - \mu\hat{N})} | r \rangle \\ &= \sum_{\{n_l\}} \exp \left[-\beta \sum_l (\varepsilon_l - \mu) n_l \right] = \sum_{\{n_l\}} \prod_l \exp [-\beta(\varepsilon_l - \mu) n_l] \\ &= \sum_{\{n_l\}} \prod_l (ze^{-\beta\varepsilon_l})^{n_l} \end{aligned}$$

where

$$\hat{H}|r\rangle = \left(\sum_l \varepsilon_l n_l \right) |r\rangle, \quad (1.40)$$

$$\hat{N}|r\rangle = \left(\sum_l \hat{n}_l \right) |r\rangle = \left(\sum_l n_l \right) |r\rangle. \quad (1.41)$$

We are now able to perform the sum in (1.34) over all the configuration independently. For bosons

$$\begin{aligned} Z_G(T, V, \mu) &= \sum_{\{n_l\}} \prod_l (ze^{-\beta\varepsilon_l})^{n_l} \\ &= \sum_{n_0=0}^{\infty} \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \dots (ze^{-\beta\varepsilon_0})^{n_0} (ze^{-\beta\varepsilon_1})^{n_1} (ze^{-\beta\varepsilon_2})^{n_2} \dots \\ &= \left[\sum_{n_0=0}^{\infty} (ze^{-\beta\varepsilon_0})^{n_0} \right] \left[\sum_{n_1=0}^{\infty} (ze^{-\beta\varepsilon_1})^{n_1} \right] \dots \\ &= \prod_{l=0}^{\infty} \sum_{n_l=0}^{\infty} (ze^{-\beta\varepsilon_l})^{n_l} = \prod_l \frac{1}{1 - e^{-\beta(\varepsilon_l - \mu)}}, \end{aligned}$$

The geometric series converges when $\mu < \varepsilon_0$. For fermions

$$Z_G(T, V, \mu) = \prod_{l=0}^{\infty} \sum_{n_l=0}^1 (ze^{-\beta\varepsilon_l})^{n_l} = \prod_l \left(1 + e^{-\beta(\varepsilon_l - \mu)} \right) \quad (1.42)$$

Equivalently we can write the grand partition function of bosons and fermions in a close form

$$Z_G = \prod_l Z_l = \prod_l [1 \mp e^{-\beta(\varepsilon_l - \mu)}]^{-1} \quad (1.43)$$

where $Z_l = [1 \mp e^{-\beta(\varepsilon_l - \mu)}]^{-1}$; $-$ is for bosons, $+$ is for fermions.

1.19 The average occupation number of ideal quantum gas

Consider a single-particle state $|l\rangle$, the average occupation number is given by

$$\begin{aligned} \langle \hat{n}_l \rangle &= \text{Tr}(\hat{\rho} \hat{n}_l) = Z_G^{-1} \text{Tr} \left[e^{-\beta(\hat{H} - \mu \hat{N})} \hat{n}_l \right] \\ &= Z_G^{-1} \sum_r \langle r | e^{-\beta(\hat{H} - \mu \hat{N})} \hat{n}_l | r \rangle \\ &= Z_G^{-1} \sum_{\{n_q\}} \exp \left[-\beta \sum_q (\varepsilon_q - \mu) n_q \right] n_l \\ &= \frac{\sum_{\{n_q\}} e^{-\beta \sum_q (\varepsilon_q - \mu) n_q} n_l}{\sum_{\{n_q\}} e^{-\beta \sum_q (\varepsilon_q - \mu) n_q}} = \frac{\Pi_q \sum_{n_q} e^{-\beta(\varepsilon_q - \mu) n_q} n_l}{\Pi_q \sum_{n_q} e^{-\beta(\varepsilon_q - \mu) n_q}} \\ &= \frac{\left(\prod_{q \neq l} Z_q \right) \sum_{n_l} e^{-\beta(\varepsilon_l - \mu) n_l} n_l}{\left(\prod_{q \neq l} Z_q \right) \sum_{n_l} e^{-\beta(\varepsilon_l - \mu) n_l}} = \frac{\sum_{n_l} e^{-\beta(\varepsilon_l - \mu) n_l} n_l}{\sum_{n_l} e^{-\beta(\varepsilon_l - \mu) n_l}} \end{aligned}$$

Therefore we can rewrite the average occupation number as the partial derivative of the sub-partition function or the grand partition function

$$\langle \hat{n}_l \rangle = \beta^{-1} \frac{\partial \ln Z_l}{\partial \mu} = -\beta^{-1} \frac{\partial \ln Z_G}{\partial \varepsilon_l} \quad (1.44)$$

$$\langle \hat{n}_l \rangle = \frac{1}{e^{\beta(\varepsilon_l - \mu)} \mp 1} \quad (1.45)$$

1.20 Maxwell-Boltzmann statistics

The canonical partition function of N Maxwell-Boltzmann particles is

$$Z_N^{\text{MB}} = \sum_{n_l=N} \frac{e^{-\beta \varepsilon_0 n_0}}{n_0!} \cdots \frac{e^{-\beta \varepsilon_l n_l}}{n_l!} \cdots \quad (1.46)$$

The $n_l!$ in the denominators can be understood as the fact that the MB particles are identical, so the state of the whole system would not be changed if permuting the n_l particles in the single-particle state with energy ε_l .

Theorem 1.2. Multinomial theorem: For any positive integer m and any nonnegative integer N , the multinomial formula tells us how a sum with m terms expands when raised to an arbitrary power N :

$$(x_1 + x_2 + \cdots + x_m)^N = \sum_{n_l=N} \binom{N}{n_1, n_2, \dots, n_m} \prod_{l=1}^m x_l^{n_l} \quad (1.47)$$

where

$$\binom{N}{n_1, n_2, \dots, n_m} = \frac{N!}{n_1! \cdots n_m!} \quad (1.48)$$

According to multinomial theorem, we can simplify the MB canonical partition function (1.46) as follows

$$\begin{aligned} Z_N^{\text{MB}} &= \sum_{n_l=N} \frac{(e^{-\beta \varepsilon_0})^{n_0}}{n_0!} \cdots \frac{(e^{-\beta \varepsilon_l})^{n_l}}{n_l!} \cdots \\ &= \frac{1}{N!} \sum_{n_l=N} \binom{N}{n_0, \dots, n_l, \dots} \prod_l (e^{-\beta \varepsilon_l})^{n_l} \\ &= \frac{1}{N!} \left(\sum_l e^{-\beta \varepsilon_l} \right)^N \\ &= \frac{(Z_1^{\text{MB}})^N}{N!} \end{aligned}$$

where Z_1^{MB} is the partition function for a single-particle system

$$Z_1^{\text{MB}} = \sum_{l=0}^{\infty} e^{-\beta \varepsilon_l} \quad (1.49)$$

Next we calculate the grand partition function

$$Z_G^{\text{MB}} = \sum_{N=0}^{\infty} z^N Z_N^{\text{MB}} = \sum_{N=0}^{\infty} z^N \frac{(Z_1^{\text{MB}})^N}{N!} = \exp(z Z_1^{\text{MB}}) \quad (1.50)$$

1.21 Riemann Zeta function

Riemann Zeta function is defined as the sum of the following infinite series

$$\zeta(\nu) = \sum_{k=1}^{\infty} \frac{1}{k^\nu}, \quad (1.51)$$

Notice that $\zeta(1)$ is known as the harmonic series. The fact that the harmonic series diverges was first proven in the 14th century by Nicole Oresme. His way to prove divergence is to compare the harmonic series with another divergent series, where each denominator is replaced with the next-largest power of two

$$\begin{aligned} & 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \dots \\ & > 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \dots \\ & = 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots \end{aligned}$$

Inspired by this comparison test, we can also consider the power of any other integer k , such that we separate the whole series into segments with $k^n - k^{n-1} = (k-1)k^{n-1}$ terms, and replace their denominator with k^n , the largest power of k in this segment, such that

$$\begin{aligned} & \frac{1}{k^{n-1}+1} + \frac{1}{k^{n-1}+2} + \dots + \frac{1}{k^n} \\ & > \frac{1}{k^n} + \frac{1}{k^n} + \dots + \frac{1}{k^n} = \frac{k-1}{k} \end{aligned}$$

There are infinite these segments in the harmonic series, so harmonic series is divergent

$$\begin{aligned} \sum_{l=1}^{\infty} &= \sum_{n=1}^{\infty} \sum_{i=1}^{(k-1)k^{n-1}} \frac{1}{k^{n-1}+i} \\ &> \sum_{n=1}^{\infty} \sum_{i=1}^{(k-1)k^{n-1}} \frac{1}{k^n} \\ &= \sum_{n=1}^{\infty} \frac{k-1}{k} = \frac{k-1}{k} \sum_{n=1}^{\infty} 1 \end{aligned}$$

Next we consider $\zeta(2)$ which is also known as the Basel problem, first posed by Pietro Mengoli in 1650 and solved by Leonhard Euler in 1734.

$$\zeta(2) = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots \quad (1.52)$$

The Basel problem had withstood the attacks of the leading mathematicians of the day, including the Bernoulli family, but none of them succeeded. This ice was broken by Euler, who was a student of Bernoulli. Euler's solution brought him immediate fame when he was twenty-eight.

Similar to our previous strategy, we can prove that this series is convergent

$$\begin{aligned}
 & 1 + \left(\frac{1}{2^2} + \frac{1}{3^2} \right) + \left(\frac{1}{4^2} + \cdots + \frac{1}{7^2} \right) + \left(\frac{1}{8^2} + \cdots + \frac{1}{15^2} \right) + \cdots \\
 & < 1 + \left(\frac{1}{2^2} + \frac{1}{2^2} \right) + \left(\frac{1}{4^2} + \cdots + \frac{1}{4^2} \right) + \left(\frac{1}{8^2} + \cdots + \frac{1}{8^2} \right) + \cdots \\
 & = 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots = \frac{1}{1 - 1/2} = 2
 \end{aligned}$$

But what is the exact value of $\zeta(2)$? Euler looked at this problem for quite a long time, and he finally figured out an amazing and creative way to solve it. His starting point is that the infinite summation can be related to the factorization of polynomial. For example, let's look at a quadratic polynomial

$$\begin{aligned}
 & \left(1 - \frac{x}{\alpha_1} \right) \left(1 - \frac{x}{\alpha_2} \right) \\
 & = 1 - ax + bx^2 \\
 & = 1 - \left(\frac{1}{\alpha_1} + \frac{1}{\alpha_2} \right) x + \frac{1}{\alpha_1 \alpha_2}
 \end{aligned}$$

where $\alpha_{1,2}$ are two roots of the polynomial $1 - ax + bx^2$, and $a = \frac{1}{\alpha_1} + \frac{1}{\alpha_2}$. By the same token, we should consider a polynomial of order $2n$ with roots $\pm 1, \pm 2, \dots, \pm n$

$$\begin{aligned}
 & \left(1 - \frac{x}{1} \right) \left(1 + \frac{x}{1} \right) \left(1 - \frac{x}{2} \right) \left(1 + \frac{x}{2} \right) \cdots \left(1 - \frac{x}{n} \right) \left(1 + \frac{x}{n} \right) \\
 & = 1 - \left(\frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2} \right) x + \cdots
 \end{aligned}$$

In order to calculate $\zeta(2)$, we need to find a polynomial whose roots lie in $\pm 1, \pm 2, \dots$. Since this polynomial is of infinite order, we should consider the Taylor expansion of some function which can represent this polynomial. According to our knowledge of the shape of functions, $\sin x$ has roots $0, \pm\pi, \pm 2\pi, \dots$, so it is natural to propose the following conjecture

$$\begin{aligned}
 \sin x &= Cx \left(1 - \frac{x}{\pi} \right) \left(1 + \frac{x}{\pi} \right) \left(1 - \frac{x}{2\pi} \right) \left(1 + \frac{x}{2\pi} \right) \cdots \\
 &= Cx \prod_{n=1}^{\infty} \left[1 - \frac{x^2}{(n\pi)^2} \right]
 \end{aligned} \tag{1.53}$$

where C is a constant which can be determined by the limit $x \rightarrow 0$

$$C = \lim_{x \rightarrow 0} \frac{\sin x}{x} = 1 \tag{1.54}$$

If conjecture (1.53) is true, we can obtain the exact value of $\zeta(2)$ by comparing the coefficient of $\sin x/x$

$$\frac{\sin x}{x} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = 1 - \frac{1}{3!}x^2 + \frac{1}{5!}x^5 + \cdots \tag{1.55}$$

Therefore

$$\frac{1}{1^2\pi^2} + \frac{1}{2^2\pi^2} + \cdots = \frac{1}{\pi^2} \left(\frac{1}{1^2} + \frac{1}{2^2} + \cdots \right) = \frac{1}{6} \tag{1.56}$$

or equivalently

$$\zeta(2) = \frac{\pi^2}{6} \tag{1.57}$$

Actually Euler's method is not restricted in calculating $\zeta(2)$, it can be generalized to calculate $\zeta(2n)$ for all even powers. To do this, we should consider a infinite order polynomial which can be factorized as an infinite series

$$1 + A_1z + A_2z^2 + A_3z^3 + \dots = (1 + \alpha_1z)(1 + \alpha_2z)(1 + \alpha_3z)\dots \quad (1.58)$$

Now we want to get the summation of every power of α_i as follows

$$\begin{aligned} S_1 &= \alpha_1 + \alpha_2 + \alpha_3 + \dots, \\ S_2 &= \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \dots, \\ S_3 &= \alpha_1^3 + \alpha_2^3 + \alpha_3^3 + \dots, \\ &\vdots \end{aligned}$$

By comparing the coefficients we obtained

$$\begin{aligned} S_1 &= A_1, \\ S_2 &= A_1S_1 - 2A_2, \\ S_3 &= A_1S_2 - A_2S_1 + 3A_3, \\ &\vdots \end{aligned} \quad (1.59)$$

Proof. Expand the infinite series as polynomial

$$\prod_{i=1}^{\infty} (1 + \alpha_i z) = 1 + \left(\sum_i \alpha_i \right) z + \left(\sum_{i < j} \alpha_i \alpha_j \right) z^2 + \left(\sum_{i < j < k} \alpha_i \alpha_j \alpha_k \right) z^3 + \dots \quad (1.60)$$

Therefore the expression of A_i are given by

$$\begin{aligned} A_1 &= \sum_i \alpha_i, \\ A_2 &= \sum_{i < j} \alpha_i \alpha_j, \\ A_3 &= \sum_{i < j < k} \alpha_i \alpha_j \alpha_k, \\ &\vdots \end{aligned}$$

Obviously $S_1 = A_1$. Next we need to prove $S_2 = A_1S_1 - 2A_2$

$$\begin{aligned} A_1S_1 &= \left(\sum_i \alpha_i \right) \left(\sum_j \alpha_j \right) \\ &= \left(\sum_{i=j} + \sum_{i \neq j} \right) \\ &= \sum_i \alpha_i^2 + 2 \sum_{i < j} \alpha_i \alpha_j \end{aligned}$$

Therefore

$$A_1 S_1 - 2A_2 = \sum_i \alpha_i^2 + 2 \sum_{i < j} \alpha_i \alpha_j - 2 \sum_{i < j} \alpha_i \alpha_j = \sum_i \alpha_i^2 = S_2 \quad (1.61)$$

For $S_3 = A_1 S_2 - A_2 S_1 + 3A_3$, we calculate the right hand side of the equation by each term

$$\begin{aligned} A_1 S_2 &= \left(\sum_i \alpha_i \right) \left(\sum_j \alpha_j^2 \right) \\ &= \left(\sum_{i=j} + \sum_{i \neq j} \right) \alpha_i \alpha_j^2 \\ &= \sum_i \alpha_i^3 + \sum_{i \neq j} \alpha_i \alpha_j^2, \end{aligned}$$

and

$$\begin{aligned} A_2 S_1 &= \left(\sum_{i < j} \alpha_i \alpha_j \right) \left(\sum_k \alpha_k \right) \\ &= \left(\sum_{i < j = k} + \sum_{k = i < j} + \sum_{i < j < k} + \sum_{i < k < j} + \sum_{k < i < j} \right) \alpha_i \alpha_j \alpha_k \\ &= \sum_{i \neq j} \alpha_i \alpha_j^2 + 3 \sum_{i < j < k} \alpha_i \alpha_j \alpha_k, \end{aligned}$$

Thus the right hand side of (1.59) is given by

$$\begin{aligned} A_1 S_2 - A_2 S_1 + 3A_3 &= \sum_i \alpha_i^3 + \sum_{i \neq j} \alpha_i \alpha_j^2 - \sum_{i \neq j} \alpha_i \alpha_j^2 - 3 \sum_{i < j < k} \alpha_i \alpha_j \alpha_k + 3 \sum_{i < j < k} \alpha_i \alpha_j \alpha_k \\ &= \sum_i \alpha_i^3 = S_3 \end{aligned}$$

The key point to calculate the above terms is changing double summations into a single summation with different constraints. \square

In our case, we choose the polynomial as $\sin x/x$ and consider the coefficients of all even powers of x , then

$$\begin{aligned} \zeta(2) &= 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = -\pi^2 S_1, \\ \zeta(4) &= 1 + \frac{1}{2^4} + \frac{1}{3^4} + \cdots = \pi^4 S_2, \\ \zeta(6) &= 1 + \frac{1}{2^6} + \frac{1}{3^6} + \cdots = -\pi^6 S_3, \\ &\vdots \\ \zeta(2n) &= (-\pi^2)^n S_n \end{aligned}$$

From (1.55), we know that $A_n = (-)^n 1/(2n+1)!$, so we can compute S_n as follows

$$\begin{aligned} S_1 &= A_1 = -\frac{1}{3!} = -\frac{1}{6} = \frac{(2\pi)^2}{2(2!)} \frac{1}{6}, \\ S_2 &= A_1 S_1 - 2A_2 = \left(-\frac{1}{3!}\right) \left(-\frac{1}{3!}\right) - 2\frac{1}{5!} = \frac{1}{90} = \frac{(2\pi)^4}{2(4!)} \frac{1}{30}, \\ S_3 &= A_1 S_2 - A_2 S_1 + 3A_3 \\ &= \left(-\frac{1}{3!}\right) \frac{1}{90} - \frac{1}{5!} \left(-\frac{1}{6}\right) + 3\left(-\frac{1}{7!}\right) \\ &= -\frac{1}{945} = \frac{(2\pi)^6}{2(6!)} \frac{1}{42} \end{aligned}$$

Therefore

$$\begin{aligned} \zeta(2) &= -\pi^2 S_1 = \frac{\pi^2}{6}, \\ \zeta(4) &= \pi^4 S_2 = \frac{\pi^4}{90}, \\ \zeta(6) &= -\pi^6 S_3 = \frac{\pi^6}{945}, \\ &\vdots \\ \zeta(2n) &= (-\pi^2)^n S_n = ? \end{aligned}$$

Later we will prove that $\zeta(2n)$ is related to Bernoulli number B_{2n} .

Theorem 1.3.

$$\zeta(2n) = (-)^{n+1} \frac{(2\pi)^{2n} B_{2n}}{2(2n)!} \quad (1.62)$$

The above results are valid if our conjecture (1.53) is true. We haven't prove that it is true! Don't worry, as one of the greatest mathematician, Euler proved this conjecture 200 years ago, or we should rewrite it as a theorem since it's been proven.

Theorem 1.4.

$$\sin x = x \prod_{n=1}^{\infty} \left[1 - \frac{x^2}{(n\pi)^2} \right] \quad (1.63)$$

Proof. It seems like we have already prove theorem 1.4 from the perspective of factorization of polynomial, but it is not rigorous because we do not know whether the infinite product converges or not. \square