Basic Elements of Statistical Mechanics

Jiasheng He   Summer 2020

draft, updated randomly

This review is a logical, self-contained short review of the basic theory of statistical mechanics. It is based on my lecture notes and reading notes of undergraduate course Physics 5600 I took in Autumn 2019 at the Ohio State University, Columbus, Ohio. The instructor of this course is Prof. Michael Poirier. The textbook we used is

D. Schroeder, *An Introduction to Thermal Physics* (Addison Wesley Longman, 2000).

Beyond the materials in this book, I also added several additional aspects which I encountered in my research projects.

Corrections and suggestions are very welcome to be sent to my e-mail address he.1326@osu.edu

---

Contents

Theory

1. Combinatorics
2. Two Einstein solids
3. Entropy and heat
4. Thermal identities
5. Tendency towards equilibrium
6. Mirco canonical, canonical, and grand canonical ensembles
7. Distribution functions

1. Combinatorics

There are $\binom{q + N - 1}{q}$ cases/ways/outcomes to distribute q units of energy to N harmonic oscillators.

For example, $q = 2$ and $N = 3$, then $\binom{2 + 3 - 1}{2} = \binom{4}{2} = 6$

Oscillator 1	Oscillator 2	Oscillator 3
2	0	0
0	2	0
0	0	2
1	1	0
1	0	1
0	1	1

2. Two Einstein solids

<u>Thermal Equilibrium</u> Two Einstein solids A and B have fixed number of harmonic oscillators N_A and N_B. How many cases/ways to distribute q units of energy to the two solids?

q_A	N_A	Ω_A	S_A	q_B	N_B	Ω_B	S_B	$\Omega = \Omega_A \Omega_B$	$S = S_A + S_B$
1	N_A			$q - 1$	N_B				
2	N_A			$q - 2$	N_B				
$\vdots$	N_A	$\vdots$	$\vdots$	$\vdots$	N_B	$\vdots$	$\vdots$	$\vdots$	$\vdots$
q_A	N_A	$\binom{q + N_A - 1}{q}$		$q - q_A$	N_B	$\binom{q - q_A + N_B - 1}{q - q_A}$		$\binom{q + N_A - 1}{q}\binom{q - q_A + N_B - 1}{q - q_A}$	

For example, the second row means if $q_A = 2$, then Ω_A and S_A are such such. It just indicates a possible case.

Stirling approximation is used at $\Omega = \Omega_A \Omega_B = \binom{q + N_A - 1}{q}\binom{q - q_A + N_B - 1}{q - q_A}$ by which the function $\Omega(U_A) = \Omega(q_A \hbar \omega)$ can be shown is a Gaussian distribution provided q and N is very large.
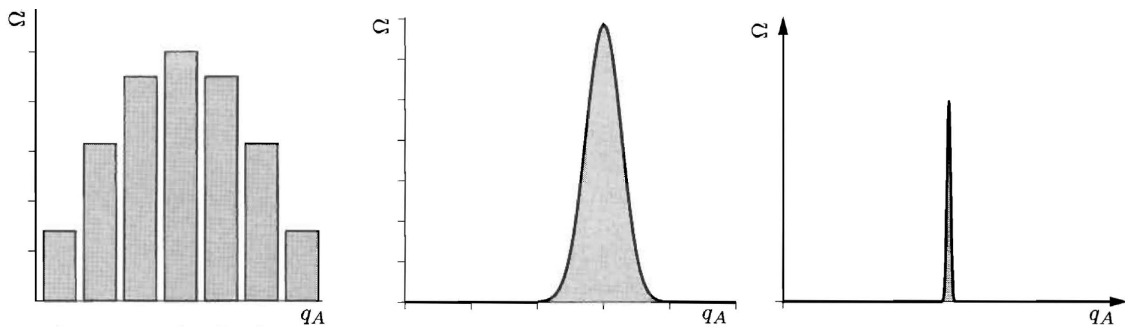

FIG. 1   Plot of total multiplicity Ω as a function of q_A.

The figure is the statistical interpretation of the second law: maximize entropy $S := k \ln \Omega$. The maximal value of S indicates the thermal equilibrium. Then we can define temperature $\dfrac{1}{T} := \left(\dfrac{\partial S}{\partial E}\right)_{V,N}$ where V and N are hold fixed. $T_A = T_B$ at thermal equilibrium.
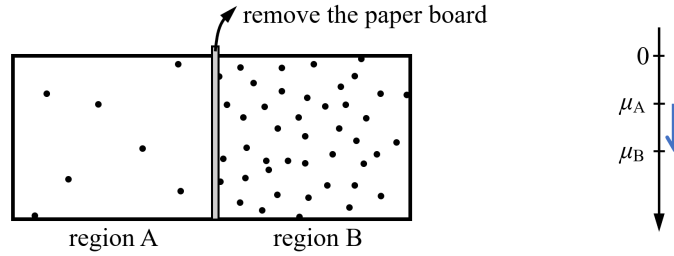
Diffusion Equilibrium

FIG. 2 (*left*) The total energy in region A is fixed. The number of particle in region A and B are not fixed yet the total number $N = N_A + N_B$ is fixed. (*right*) The particles tend to move to the region with smaller chemical potential.

q_A	N_A	Ω_A	S_A	q_B	N_B	Ω_B	S_B	$\Omega = \Omega_A \Omega_B$	S
q_A	1			q_B	$N-1$				
q_A	2			q_B	$N-2$				
q_A	$\vdots$	$\vdots$	$\vdots$	q_B	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
q_A	N_A	$\binom{q+N_A-1}{q}$		q_B	$N-N_A$	$\binom{q_B+N-N_A-1}{q_B}$		$\binom{q+N_A-1}{q}\binom{q_B+N-N_A-1}{q_B}$	

Again, the plot shows The maximal value of S indicates the diffusive equilibrium. At diffusive equilibrium, $0 = \dfrac{\partial S}{\partial N_A} = \dfrac{\partial S_A}{\partial N_A} + \dfrac{\partial S_B}{\partial N_A} = \dfrac{\partial S_A}{\partial N_A} - \dfrac{\partial S_B}{\partial N_B}$. So, we define the chemical potential $\dfrac{\mu}{T} := -\left(\dfrac{\partial S}{\partial N}\right)_{E,V}$ provided E and V fixed. At diffusive equilibrium, $\mu_A = \mu_B$.

Note that $\mu \leqslant 0$. See Figure 2 (right). $0 > \mu_A > \mu_B$ means region B tends to gain particles.
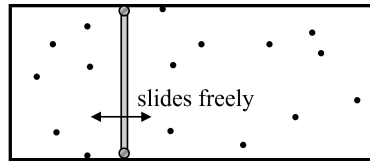
Mechanical Equilibrium

FIG. 3   The board can move frictionlessly due to pressure difference on both sides.

Similarly, in this case, we can define pressure $\dfrac{P}{T} := \left(\dfrac{\partial S}{\partial V}\right)_{E,N}$ where E and N are fixed.

3. Entropy and heat From $\dfrac{1}{T} = \left(\dfrac{\partial S}{\partial E}\right)_{V,N}$, we have if V, N are fixed, the work $W = 0$, and T does not vary in system, then $\mathrm{d}S = \dfrac{\mathrm{d}E}{T} = \dfrac{Q}{T}$.

4. Thermal identities

Based on the discussions in 2., we can say that entropy S is a function of E, V, N in general. Then

$$\mathrm{d}S = \frac{\partial S}{\partial E}\mathrm{d}E + \frac{\partial S}{\partial V}\mathrm{d}V + \frac{\partial S}{\partial N}\mathrm{d}N = \frac{1}{T}\mathrm{d}E + \frac{P}{T}\mathrm{d}V - \frac{\mu}{T}\mathrm{d}N$$

So, we have

$$\mathrm{d}E = T\mathrm{d}S - P\mathrm{d}V + \mu\mathrm{d}N \tag{1}$$

This means E is a function of S, V, and N.

Definition of Legendre transformation The Legendre transformation of a convex function $f : X \to \mathbb{R}$ defined on a convex set $X \subseteq \mathbb{R}^n$ is the function $f^* : X^* \to \mathbb{R}$

$$f^*(x^*) = \sup_{x \in X} (f(x) - \langle x^*|x\rangle)$$

where $X^* = \left\{x^* \in \mathbb{R}^n : \sup_{x \in X} (f(x) - \langle x^*|x\rangle) < \infty\right\}$.

Our definition is a little bit different from the standard definition in any textbook of convex analysis. In the better and loose cases, the convex set is an interval of $\mathbb{R}$ and $f^*(x^*) = \sup_{x \in X}(f(x) - x^*x)$. To find the "sup" with respect to $x \in X$, we want $0 = \dfrac{\mathrm{d}f^*}{\mathrm{d}x} = \dfrac{\mathrm{d}f}{\mathrm{d}x} - x^*$, so

$$x^* = \frac{\mathrm{d}f}{\mathrm{d}x}$$

Note that $\dfrac{\mathrm{d}f^*}{\mathrm{d}x^*} = -x$.

Enthalpy $H(S, P, N)$: Use $-P = \dfrac{\partial E}{\partial V}$ to perform Legendre transformation on $E(S, V, N)$, we can get enthalpy $H = E - (-P)V = E + PV$ and thus

$$\mathrm{d}H = T\mathrm{d}S + V\mathrm{d}P + \mu\mathrm{d}N \tag{2}$$

Note that $-V = \dfrac{\partial H}{\partial(-P)}$. The physical meaning of enthalpy is the energy required to exist in an environment with pressure P.

Helmholtz Free Energy $F(T, V, N)$: Use $T = \dfrac{\partial E}{\partial S}$ to perform Legendre transformation on $E(S, V, N)$, we can get Helmholtz free energy $F = E - TS$ and thus

$$\mathrm{d}F = \mathrm{d}E - T\mathrm{d}S - S\mathrm{d}T = -S\mathrm{d}T - P\mathrm{d}V + \mu\mathrm{d}N \tag{3}$$

Note that $-S = \dfrac{\partial F}{\partial T}$. $E = TS + F$ indicates that the physical meaning of Helmholtz free energy is the net energy you need to input for a substance to exist in an environment (pressure is not considered). Note that TS is the automatic input from the environment.

Gibbs Free Energy $G(T, P, N)$: Use $T = \dfrac{\partial H}{\partial S}$ to perform Legendre transformation on $H(S, P, N)$, we can get Gibbs free energy $G = H - TS$ and thus

$$\mathrm{d}G = \mathrm{d}H - T\mathrm{d}S - S\mathrm{d}T = -S\mathrm{d}T + V\mathrm{d}P + \mu\mathrm{d}N \tag{4}$$

Note that $-S = \dfrac{\partial G}{\partial T}$. Gibbs free energy $G(T, P, N)$ can also be transformed from $F(T, V, N)$ by using $-P = \dfrac{\partial F}{\partial V}$. $H = E + PV = G + TS$ indicates that the physical meaning of Gibbs free energy is the net energy you need to input for a substance to exist in an environment with pressure P. Note that TS is the automatic input from the environment.

Grand Free Energy $\Phi(T, V, \mu)$: Use $\mu = \dfrac{\partial F}{\partial N}$ perform Legendre transformation on $F(T, V, N)$, we can get grand potential energy $\Phi = F - \mu N$ and thus

$$\mathrm{d}\Phi = \mathrm{d}F - \mu\mathrm{d}N - N\mathrm{d}\mu = -S\mathrm{d}T - P\mathrm{d}V - N\mathrm{d}\mu \tag{5}$$

Note that $-N = \dfrac{\partial \Phi}{\partial \mu}$. $E = \Phi + TS + \mu N$ indicates that the physical meaning of grand free energy is the net energy you need to input for a substance to exist in an environment (pressure is not considered). Note that TS and μN are the automatic inputs from the environment; in other words, the system is allowed only to exchange energy and particle with the environment.
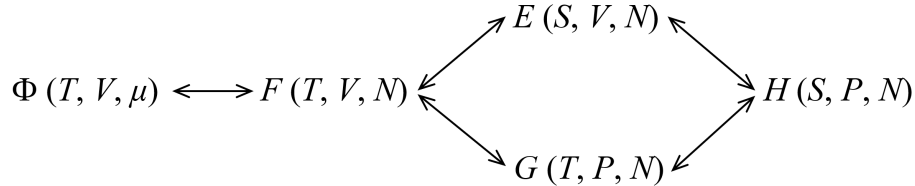
A summary picture is given

FIG. 4   Thermal quantities are related by Legendre transformation.

5. Tendency towards equilibrium Consider a system contacting with a reservoir.

Assumption 1: $T = T_{\mathrm{rsv}}$, N_{rsv}, and V_{rsv} are fixed. Then,

$$\mathrm{d}S_{\mathrm{total}} = \mathrm{d}S + \mathrm{d}S_{\mathrm{rsv}} = \left(\frac{1}{T}\mathrm{d}U + \frac{P}{T}\mathrm{d}V - \frac{\mu}{T}\mathrm{d}N\right) + \frac{1}{T}\mathrm{d}U_{\mathrm{rsv}} \xlongequal{\mathrm{d}U=-\mathrm{d}U_{\mathrm{rsv}}} \frac{1}{T}\left(P\mathrm{d}V - \mu\mathrm{d}N\right) = -\frac{1}{T}\mathrm{d}F,$$

by using the second law saying S tends to increase, we have

Conclusion 1: If $T = T_{\mathrm{rsv}}$, N_{rsv}, and V_{rsv} are fixed, then F tends to decrease.

Assumption 2: $T = T_{\mathrm{rsv}}$, N_{rsv}, and $P = P_{\mathrm{rsv}}$ are fixed, but V_{rsv} is not fixed. Then,

$$\mathrm{d}S_{\mathrm{total}} = \mathrm{d}S + \mathrm{d}S_{\mathrm{rsv}} = \left(\frac{1}{T}\mathrm{d}U + \frac{P}{T}\mathrm{d}V - \frac{\mu}{T}\mathrm{d}N\right) + \left(\frac{1}{T}\mathrm{d}U_{\mathrm{rsv}} + \frac{P}{T}\mathrm{d}V\right) \xlongequal[\mathrm{d}V=-\mathrm{d}V_{\mathrm{rsv}}]{\mathrm{d}U=-\mathrm{d}U_{\mathrm{rsv}}} -\frac{1}{T}\mu\mathrm{d}N = -\frac{1}{T}\mathrm{d}G,$$

by using the second law saying S tends to increase, we have

Conclusion 2: If $T = T_{\mathrm{rsv}}$, N_{rsv}, and $P = P_{\mathrm{rsv}}$ are fixed, then G tends to decrease.

6. Micro canonical ensemble: the system cannot exchange anything with the environment.
Canonical ensemble: the system can only exchange energy with the environment
Grand canonical ensemble: the system can only exchange energy and particles with the environment.
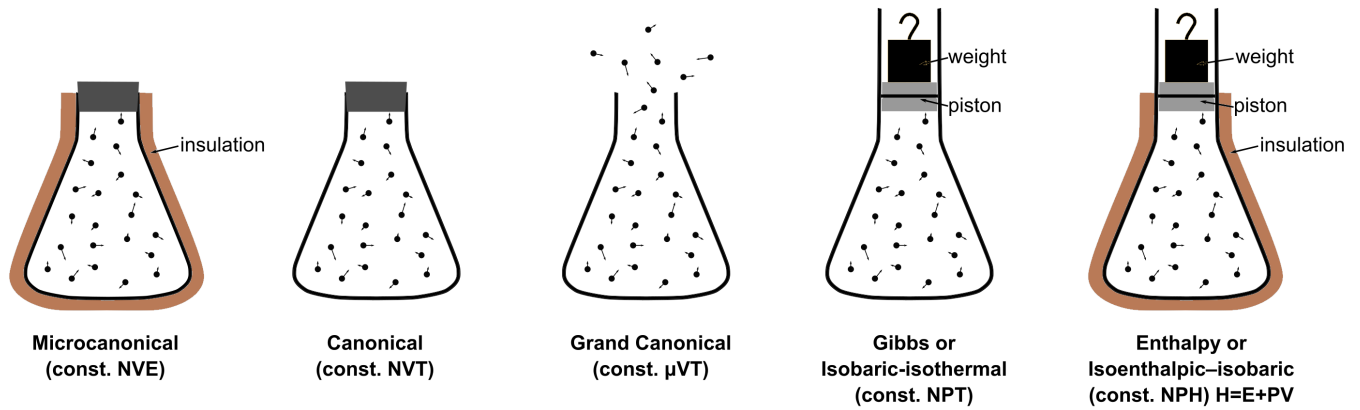

FIG. 5 Statistical Ensembles. The relations between system and the environment. The image is from `wikipedia.org`

Micro Canonical Ensemble
?????

Canonical Ensemble
?????

Grand Canonical Ensemble *using grand partition function $\mathcal{Z}$*

The system is in contact with a reservoir; it is at thermal equilibrium with a reservoir and can only exchange energy and particles with the reservoir.

Again, we have $\dfrac{\text{Prob}(s_2)}{\text{Prob}(s_1)} = \dfrac{\Omega_{\text{rsv}}(s_2)}{\Omega_{\text{rsv}}(s_1)} = \dfrac{\exp\left(\dfrac{S_{\text{rsv}}(s_2)}{k}\right)}{\exp\left(\dfrac{S_{\text{rsv}}(s_1)}{k}\right)} = e^{\frac{1}{k}(S_{\text{rsv}}(s_2)-S_{\text{rsv}}(s_1))}$.

Invoke $\mathrm{d}S_{\text{rsv}} = \dfrac{1}{T_{\text{rsv}}}(\mathrm{d}U_{\text{rsv}} + P\,\mathrm{d}V_{\text{rsv}} - \mu\,\mathrm{d}N_{\text{rsv}}) = \dfrac{1}{T_{\text{rsv}}}(\mathrm{d}U_{\text{rsv}} - \mu\,\mathrm{d}N_{\text{rsv}})$ where we set $\mathrm{d}V_{\text{rsv}} = 0$ only. This is Gibbs statistics.

Because of thermal equilibrium, $T_{\text{sys}} = T$. We have

$$S_{\text{rsv}}(s_2) - S_{\text{rsv}}(s_1) = \frac{1}{T}\big(E_{\text{rsv}}(s_2) - E_{\text{rsv}}(s_1) - \mu(N_{\text{rsv}}(s_2) - N_{\text{rsv}}(s_1))\big)$$

To change from reservoir to system, use conservation of energy $E_{\text{rsv}}(s_2) + E(s_2) = E_{\text{rsv}}(s_1) + E(s_1)$ and of particles $N_{\text{rsv}}(s_2) + N(s_2) = N_{\text{rsv}}(s_1) + N(s_1)$, we have

$$S_{\text{rsv}}(s_2) - S_{\text{rsv}}(s_1) = \frac{1}{T}\big(E(s_1) - E(s_2) - \mu(N(s_1) - N(s_2))\big)$$

$$= -\frac{1}{T}\big((E(s_2) - \mu N(s_2)) - (E(s_1) - \mu N(s_1))\big)$$

So, $\dfrac{\text{Prob}(s_2)}{\text{Prob}(s_1)} = e^{-\frac{1}{kT}\left((E(s_2)-\mu N(s_2))-(E(s_1)-\mu N(s_1))\right)} = \dfrac{e^{-\frac{1}{kT}(E(s_2)-\mu N(s_2))}}{e^{-\frac{1}{kT}(E(s_1)-\mu N(s_1))}}$; the ratio of $\text{Prob}(s)$ is the ratio of $e^{-\frac{E(s)-\mu N(s)}{kT}}$.

Jiasheng He, 2020

Gibbs factor is $e^{-\frac{E(s)-\mu N(s)}{kT}}$ for a microstate s of the system.

In general, for many different species of particles, Gibbs factor is $e^{-\frac{E(s)-\sum_i \mu_i N_i(s)}{kT}}$.

Gibbs probability $\text{Prob}(s) = \frac{1}{\mathcal{Z}} e^{-\frac{E(s)-\mu N(s)}{kT}}$ where normalization gives

$$\mathcal{Z} = \sum_s e^{-\frac{E(s)-\mu N(s)}{kT}} \text{ , the grand partition function}$$

$$= \sum_j e^{-\frac{E(s_j)-\mu N(s_j)}{kT}}$$

$$= \sum_j e^{-\frac{\langle s_j|H|s_j\rangle - \mu \langle s_j|\hat{N}|s_j\rangle}{kT}}$$

$$= \sum_j e^{-\frac{H_{jj}-\mu \hat{N}_{jj}}{kT}}$$

Note that $\hat{H}$ and $\hat{N}$, written in the basis of microstates of the system s_j, are diagonal matrices; so is the matrix $\hat{H} - \mu \hat{N}$. Hence, by

$$e^{-\frac{\hat{H}-\mu\hat{N}}{kT}} = \begin{pmatrix} e^{-\frac{\hat{H}_{11}-\mu\hat{N}_{11}}{kT}} & & \\ & e^{-\frac{\hat{H}_{22}-\mu\hat{N}_{22}}{kT}} & \\ & & \ddots \end{pmatrix} = \begin{pmatrix} e^{-\frac{E(s_1)-\mu N(s_1)}{kT}} & & \\ & e^{-\frac{E(s_2)-\mu N(s_2)}{kT}} & \\ & & \ddots \end{pmatrix}$$

we have grand partition function $\mathcal{Z} = \text{tr}\left(e^{-\frac{\hat{H}-\mu\hat{N}}{kT}}\right)$.

Theorem For grand canonical ensemble (Gibbs statistics at equilibrium), T and V are fixed, then we have $\Phi = -kT \ln \mathcal{Z}$.

In above discussion, the letter s indicates the state of the *whole* system. The system is made up by several same-species particles, and each particle is in **a single-particle state f_i which is the tensor product of the linear combination of spatial eigen-solutions ψ_0, ψ_1, ψ_2, ... and the linear combination of spin solutions, of single-particle Schröedinger equation**. So, the system state s can be written as an ordered pair of single-particle states f_i. For example, suppose a helium gas is composed of three helium atoms, then $s = (f_1, f_2, f_3)$. From this point of view, the energy $E(s)$ is just the sum of the energy of each single-particle state if each is stationary, but $N(s)$ is hard to tell due to exchanging with reservoir. However, we have the following logic.

A specific single-particle state f_i with energy ε_i can be occupied by 0, 1, 2, ..., or more same-species particles; so, the grand partition function of f_i is $\sum_{n=0}^{\text{all}} e^{-\frac{n(\varepsilon_i - \mu)}{kT}}$. Hence, the total grand partition function (multiplicity of the system) is their product $\prod_i \sum_{n=0}^{\text{all}} e^{-\frac{n(\varepsilon_i - \mu)}{kT}}$.

> Let E and N be the energy and particle's number of the system state s, and let ε and n be the energy and particle's number of a single-particle state f, then the grand partition function of the system is
>
> $$\mathcal{Z} = e^{-\frac{\Phi}{kT}} = \sum_s e^{-\frac{E(s)-\mu N(s)}{kT}} = \text{tr}\left(e^{-\frac{\hat{H}-\mu \hat{N}}{kT}}\right) = \prod_\varepsilon \sum_{n=0}^{\text{all}} e^{-\frac{n(\varepsilon-\mu)}{kT}} \tag{6}$$

Grand Canonical Ensemble *using grand free energy* Φ

By the preceding theorem, we can rewrite the probability of the system in state s as $\text{Prob}(s) = e^{-\frac{E-\Phi-\mu N}{kT}}$, Gibbs probability.

7. Distribution functions Let the energy of a single-particle state be ε, then the probability of n such particles in this state is

$$\text{Prob}(n \text{ in state } s) = \frac{1}{\mathcal{Z}} e^{-\frac{n\varepsilon - \mu n}{kT}} = \frac{1}{\mathcal{Z}} e^{-\frac{n(\varepsilon - \mu)}{kT}}$$

<u>Fermi-Dirac Distribution</u> If n particles are n fermions, then $n = 0$ or 1, then

$$\mathcal{Z} = e^0 + e^{-\frac{\varepsilon - \mu}{kT}} = 1 + e^{-\frac{\varepsilon - \mu}{kT}}$$

$$\langle n \rangle = \sum_{n=0}^{1} n \, \text{Prob}(n \text{ in state } s) = 0 + 1 \cdot \text{Prob}(1 \text{ in state } s) = 0 + \frac{1}{1 + e^{-\frac{\varepsilon - \mu}{kT}}} e^{-\frac{\varepsilon - \mu}{kT}} = \frac{1}{e^{\frac{\varepsilon - \mu}{kT}} + 1}$$

> **Fermi-Dirac Distribution**
> Average number of fermions in a single-particle state (energy ε) is $\langle n \rangle_{\text{FD}} = \dfrac{1}{e^{\frac{\varepsilon - \mu}{kT}} + 1}$.

<u>Bose-Einstein Distribution</u> If n particles are n bosons, then n can be any non-negative integers.

$$\mathcal{Z} = 1 + e^{-\frac{\varepsilon - \mu}{kT}} + e^{-\frac{2(\varepsilon - \mu)}{kT}} + \cdots = \sum_{n=0}^{\infty} e^{-nx} = \frac{1}{1 - e^{-x}} \text{ where } x = \frac{\varepsilon - \mu}{kT}$$

$$\langle n \rangle = \sum_{n=0}^{\infty} n \frac{1}{\mathcal{Z}} e^{-nx} = -\frac{1}{\mathcal{Z}} \sum_{n=0}^{\infty} (-n) e^{-nx} = -\frac{1}{\mathcal{Z}} \sum_{n=0}^{\infty} \frac{\partial}{\partial x}(e^{-nx}) = -\frac{1}{\mathcal{Z}} \frac{\partial}{\partial x} \left(\sum_{n=0}^{\infty} e^{-nx} \right) \text{ since the series converges}$$

$$= -\frac{1}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial x} = \text{(plug in } \mathcal{Z} \text{ and do some basic calculations)} = \frac{1}{e^x - 1}$$

> **Bose-Einstein Distribution**
> Average number of bosons in a single-particle state (energy ε) is $\langle n \rangle_{\text{BE}} = \dfrac{1}{e^{\frac{\varepsilon - \mu}{kT}} - 1}$.

<u>Boltzmann Distribution</u> Boltzmann statistics tells us that the probability of a particle being in the state (energy ε) is $\dfrac{1}{Z_{\text{single}}} e^{-\frac{\varepsilon}{kT}}$. So if we have N such particles (a general speaking, so we will not specify whether they are a kind of fermion or a kind of boson), then the average number of particles in this state is

$$N \left(\frac{1}{Z_{\text{single}}} e^{-\frac{\varepsilon}{kT}} \right) \xlongequal{\mu = -kT \ln \frac{Z_{\text{single}}}{N}} e^{\frac{\mu}{kT}} e^{-\frac{\varepsilon}{kT}} = e^{-\frac{\varepsilon - \mu}{kT}}$$

> **Boltzmann Distribution**
> Average number of particles in a single-particle state (energy ε) is $\langle n \rangle_{\text{Boltzmann}} = \dfrac{1}{e^{\frac{\varepsilon - \mu}{kT}}}$.