Lecture 1

Topics of Simulations

- Molecule Dynamics (may not be in equilibrium)
- Monte Carlo Simulations (must be in equilibrium)

What is Molecule Dynamics?

Molecule Dynamics (MD), is a subject to study the interaction within many-particles system at atomic resolutions. Reasons:

1. Find out equation of state
2. Get continuum properties such as boundary conditions, transport coefficients ···

MD (or MC) aims to build up the connection between microscopic information and macroscopic properties of physics system. However, the normal number of atoms of typical macro systems in daily life are order about $10^{23}$, while the current simulations could do $10^9$, at maximum, with the help of most advanced computer we have, which is much smaller than the real system. Although the big difference of order, we could still extract useful information from computer. That’s mainly owe to the **Sterling’s Approximation**, which is quite a good approximation for even $n > 10$ in statistical mechanics.

$$n! = \sqrt{2\pi n} e^{n \log n - n}$$

Stirling’s Formula

Introduction to Basic Interactions

From physics, we know the following fundamental interactions between atoms in MD: Coulomb, Magnetism and Gravitation as described below.

$$V_{coul} = \frac{q_1 q_2}{4\pi \varepsilon_0 r} \approx 10^{-18} J$$

$$V_{mag} = \left(\frac{v}{c}\right)^2 V_{coul} \longrightarrow 0$$

$$V_{grav} = mg \Delta h \approx 10^{-27} (\text{kg}) \cdot 10(\text{m/s}^2) \cdot 10^{-6} \text{m} \approx 10^{-32} \text{J} \longrightarrow 0$$

Here comes a fundamental problem in simulation. For a system of $N$ atoms, there will be approximately $N^2$ interaction, which grows quickly when we have a big $N$ system. This challenge principally forbidden us away from a large system simulation. We have kept searching good ways to overcome this short bar. Fortunately, screening helps a lot, e.g:

$$V(\vec{r}) = \sum_i \frac{q_i}{4\pi \varepsilon_0 |\vec{r} - \vec{r_i}|} = \frac{1}{4\pi \varepsilon_0} \left( \frac{Q}{r} + \frac{\vec{p} \cdot \vec{r}}{r^3} + \cdots \right)$$
Let’s study the general range of interaction

\[
E = \int_{\sigma}^{R} d^3 \rho(r) \frac{q}{4\pi\epsilon_0 r^n} = \int_{\sigma}^{R} dr \rho \left( \frac{k}{r^{n-2}} \left\{ \begin{array}{ll}
R^{3-n} - \sigma^{3-n} & \text{if } n \neq 3 \\
\frac{3-n}{\log \left( \frac{R}{\sigma} \right)} & \text{if } n = 3
\end{array} \right. \right)
\]

From the above simple result, we know that for \( R \to \infty \), the energy is finite for \( n > 3 \), while divergent for \( n \leq 3 \). So we would have a short-range interaction and a \( O(N) \) interaction after screening only when we have a \( n > 3 \) interaction. That’s the key point we should remember when we say whether the interactions are short-range or not.

Actually, one widely used interaction between neutral molecules or atoms is called Lennard-Jones, which was first proposed in 1924 by John Lennard-Jones. We will fully discuss it on next lecture. Before that we give a short proof from the view of quantum mechanics.

**Interaction between Two Harmonic Oscillators**

\[
V_{coul} = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r_{ij}}
\]

\[
V_{d-d} = \frac{1}{4\pi\epsilon_0} \frac{1}{r^3} \left[ 3(\vec{p}_1 \cdot \vec{r})(\vec{p}_2 \cdot \vec{r}) - \vec{p}_1 \cdot \vec{p}_2 \right]
\]

\[
V_{LJ} = 4\epsilon \left[ \left( \frac{r_{ij}}{\sigma} \right)^{-12} - \left( \frac{r_{ij}}{\sigma} \right)^{-6} \right]
\]

**Lennard Jones Potential**

The last interaction above is so called Lennard-Jones potential, which is the most common interaction in simulations. There are two parts in it: the second term with \( r^{-6} \) corresponds to the attractive force, often called van der Waals Force or Dispersion Force, which is caused by instantaneous induced dipole-dipole forces The first term with \( r^{-12} \) gives rise the Pauli repulsion, the form of which is chosen for convenience (actually it could be chosen other forms such as \( e^{-r} \)). The variable appeared in LJ potential are respectively:

- \( \epsilon \): energy depth, which governs the strength of interaction,
• $\sigma$: distance, at which the potential is zero,

• $r_{ij} = |\vec{r}_i - \vec{r}_j|$: distance between the particles.

We fit these parameter to experiments get the numerical number fro them. For example, element Argon, $\epsilon/k_B = 120$K, $\sigma = 3.4\text{Å}$.

![Figure 1: Lennard-Jones Potential](image)

To apply the idea of screening, we would cut-off the original Lennard-Jones Potential, first let’s just truncate it at $r_c$, (Typically $r_c = 2.5\sigma$)

$$V_1(r_{ij}) = \begin{cases} V_{LJ}(r) & r_{ij} < r_c \\ 0 & r_{ij} \geq r_c \end{cases}$$

Then, to overcome the discontinuity in potential and to keep the conservation of energy, we shift it by a constant

$$V_2(r_{ij}) = \begin{cases} V_{LJ}(r) - V_{LJ}(r_c) & r_{ij} < r_c \\ 0 & r_{ij} \geq r_c \end{cases}$$

Finally, let’s rule out the discontinuity in force acceleration so as to get a more smooth potentail. The final modified potential being used in practice is:

$$V_3(r_{ij}) = \begin{cases} V_{LJ}(r) - V_{LJ}(r_c) - (r - r_c)V'_{LJ}(r_c) & r_{ij} < r_c \\ 0 & r_{ij} \geq r_c \end{cases}$$
MD units

In MD, we employ $\sigma, m$ and $\epsilon$ as our fundamental quantities, so any other quantities could be transferred i.e.:
- length: $r \rightarrow r^* = \frac{r}{\sigma}$
- energy: $e \rightarrow e^* = \frac{e}{\epsilon}$
- time: $t \rightarrow t^* = \frac{t}{\sqrt{m\sigma^2/\epsilon}}$
- gravity: $g \rightarrow g^* = \frac{g}{\epsilon/m}$

Lecture 3

by Prof. Illona Kretzschmar

MD Simulations with empirical models

- Born-Oppenheimer assumption: $e^-$ move much faster than nucleus
- Treat the nuclei as point particles that follow classical Newtonian dynamics
- System stays in ground state through simulation

Assume $N$ nuclei and $Ne$ electron, and denote dynamic variables as

Coordinates: $\vec{R}_1, \ldots, \vec{R}_N, \vec{r}_1, \ldots, \vec{r}_N$
Momentum: $\vec{P}_1, \ldots, \vec{P}_N, \vec{p}_1, \ldots, \vec{p}_N$
Mass: $M_1, \ldots, M_N, m_1, \ldots, m_N$

Hamiltonian of the dynamic system:

$$H = \sum_{i=1}^{N} \frac{P_i^2}{2m_i} + \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + \sum_{i<j}^{N} \frac{e^2}{|\vec{r}_i - \vec{r}_j|} + \sum_{i<j} \frac{z_i^2 e^2}{|\vec{R}_i - \vec{R}_j|} - \sum_{i,I} \frac{z_I e^2}{|\vec{R}_I - \vec{r}_i|}$$

Schrödinger Equation (time independent)

$$[T_N + T_E + V_{ee} + V_{NN} + V_{eN}] \Psi(x, R) = E \Psi(x, R)$$

in which $x =$ (position of $e^-$ & spin variable)

Separation Ansatz (Starting point)

$$\Psi(x, R) = \Phi(x, R) \chi(R)$$
\[ \chi(R) = \text{nuclei wave function} \]

\[ \Phi(x, R) = \text{electron wave function dependent on the positions of nuclei} \]

Then we get one equation with different variable on two sides

\[ \frac{[T_e + V_{ee} + V_{eN}] \Psi(x, R)}{\Phi(x, R)} = \frac{[E - T_N - V_{NN}] \chi(R)}{\chi(R)} = \text{Const}. \]

The first term above give the electrons’ Hamiltonian and Shrödinger Equation for electrons:

\[ H_e(R) = T_e + V_{ee} + V_{eN} \]

\[ H_e(R) \Phi(x, R) = \epsilon(R) \Phi(x, R) \]

We call \( \epsilon(R) \) Potential Energy Surf or Born-Oppenheimer Surface.

There are two methods to solve the above equation:

- Density Function Theory (DFT)
- Quantum Chemical ab initio MD

**Ab Initio MD**

proposed by Car and Parrinello in 1985. ⋯

### Lecture 4

**Pseudo-Random Number**

It is called linear congruential generator algorithm[3, 4], which is easy and fast. First job is to make some large integers such as \( a, c, m \), also one integer called seed \( I_0 \). The generator is defined by recurrence formula:

\[ I_{j+1} = (aI_j + c) \mod (m) \]

There are several points to remind:

1. \( m - 1 = \text{largest possible integer} = \text{largest number that fit} \), for 32 bits OS, \( m = 2^{32} - 1 \);
2. not all \( \{a, c, m\} \) work, you need to check that \( \{I_j/m\} \in (0, 1) \) are uniformly populated.

   Actually, we could do further test, e.g.: group \( \{I_j\} \) in pairs such as \( \{(x_i, y_i)\} \) or even better in triplets \( \{(x_i, y_i, z_i)\} \). And then by checking the correlation within these points, we would have a idea whether this generator are good or not.
Monte Carlo Integration

To do an integration such as

\[
\int_{a}^{b} d x f(x) = (b - a) \int_{a}^{b} d x \frac{1}{b - a} f(x) = (b - a) \langle f \rangle_{p(x)}
\]

In which:

\[
\langle f \rangle_{p(x)} = \int_{a}^{b} d x p(x) f(x) \quad p(x) = \frac{1}{b - a}
\]

We would choose a series of \( \{x_i\} \) with pdf \( p(x) \) on domain \((a, b)\), then the original integration is done by summation and averaging

\[
\langle f \rangle_{p(x)} = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} f(x_i)
\]

This is simple MC integration, mathematically

\[
\int_{a}^{b} d x f(x) = (b - a) \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} f(x_i)
\]

Practically

\[
I_{MC} = \frac{b - a}{N} \sum_{i=1}^{N} f_i \quad f_i = f(x_i)
\]

To understand the accuracy or convergence performance of MC, we would calculate the standard deviation:

\[
\sigma^2 = \langle (I - \langle I \rangle)^2 \rangle = \langle I^2 \rangle - \langle I \rangle^2
\]

\[
= \left( \frac{b - a}{N} \right)^2 \left[ \sum_{i,j} \langle f_i f_j \rangle - \sum_{i} \langle f_i \rangle \sum_{j} \langle f_j \rangle \right]
\]

we denote

\[
\langle f_i \rangle = \bar{f}
\]

\[
\langle f_i f_j \rangle = \begin{cases} \bar{f}^2 & i = j \\ \bar{f}^2 & i \neq j \end{cases}
\]

Therefore

\[
\sigma^2 = \left( \frac{b - a}{N} \right)^2 \left[ N\bar{f}^2 + N(N - 1)\bar{f}^2 - N^2\bar{f}^2 \right]
\]

\[
= \frac{(b - a)^2}{N} \left[ \bar{f}^2 - \bar{f}^2 \right] \sim N^{-1}
\]
We just showed that the standard deviation decays as $N^{-1}$. While as we know in Non-MC Method for numerical integration, the error decay as $h^k \sim N^{-k/d}$. If $d > k$, it will be less efficient than MC integration. That’s the main reason why we need MC.

Here is an example to show the big problem you would find in MC: slow convergence and doesn’t always work.

$$\int_0^1 d \ x \ 1000e^{-1000x} = 1 - e^{-1000} \approx 1$$

It could be hard to get correct convergence with small number of MC sampling points. The problem lies in that this integrated function varying widely, which cut down the efficiency of the sampling points. The way to get rid of this weakness is often called Importance Sampling MC. Let’s rewrite the integration as

$$\int_a^b d \ x \ f(x) = \int_a^b d \ x \ p(x) \frac{f(x)}{p(x)}$$

$$= \int_a^b d \ x \ – p(x) \frac{f(x)}{p(x)} \int_a^b p(x) \ dx$$

$$= \left\langle \frac{f(x)}{p(x)} \right\rangle_{\rho(x)} \int_a^b p(x) \ dx$$

in which

$$\left\langle \frac{f(x)}{p(x)} \right\rangle_{\rho(x)} = \int_a^b d \ x \ \rho(x) \frac{f(x)}{p(x)}$$

If we could generate sampling points $\{x_i\}$ with pdf $\rho(x)$, then the integration will be just summations and averaging:

$$I_{MC} = \frac{1}{N} \sum_{i=1}^N \frac{f_i}{p_i}$$

The benefit of this procedure is that we could choose $p(x)$ such that $\frac{f(x)}{p(x)}$ varies slowly. Or say that $p(x)$ should be a decent function which looks like $f(x)$. Thus it will speed up the convergence or the efficient rate of sampling. To understand it mathematically, the standard deviation would be:

$$\sigma^2 = \frac{1}{N} \left[ \left( \frac{f}{p} \right)^2 - \left( \frac{f}{p} \right)^2 \right] \sim N^{-1}$$

Although still goes as $N^{-1}$, it could be dramatically reduced by choosing a $p(x)$ to make $(f(x)/p(x))$ more stable than $f(x)$ alone. So the left job is to generate points with required pdf, which will be covered next time.
Lecture 5

Mathematics of Probability

In a Simple MC integration, we need a uniform distribution random numbers. To improve the efficiency, we adopt importance sampling, requiring probabilities numbers of a specific distribution. Let $X = (\text{random variable with value } x)$, and define:

- $p(x) = \text{probability distribution function (pdf)}$.
- $c(x) = \text{accumulate distribution function} = \int_{-\infty}^{x} d'x'p(x') = \text{Prob}(X < x)$

with the following properties: (1) $c'(x) = p(x) > 0$, (2) $c(-\infty) = 0, c(+\infty) = 1$, (3) $c(x) \leftrightarrow p(x)$ is a one-one correspondence analytically or numerically. Next question is how to generate $x_i$ with pdf $p(x)$? We follow just two step as,

Step 1: generate a series of $y_i$ in random uniform $\in (0, 1)$

Step 2: inverse it to $x_i = c^{-1}(y_i)$

Claim: the set of $x_i$ have pdf $p(x)$

Proof 1:

$$\text{Prob}(x_i < x) = \text{Prob}(c^{-1}(y_i) < x) = \text{Prob}(y_i < c(x)) = c(x)$$

Proof 2:

$$P_X \, dx = P_Y \, dy \quad \text{with} \quad y = c(x)$$

$$\Rightarrow$$

$$P_X = P_Y \frac{dy}{dx}$$

$$= P_Y \cdot c'(x)$$

$$= 1 \cdot P_X(x)$$

Example: $p(x) = \frac{1}{a}e^{-ax} \quad (0 \leq x \leq \infty)$ with normalization $\int_{0}^{\infty} \, dx = 1$, and accumulate function $c(x) = \int_{0}^{x} \, dx'p(x')$. Then we choose $y_i \in (0, 1)$ uniformly, then get the corresponding $x_i = c^{-1}(y_i) = -\frac{1}{a} \log(1 - y_i)$, which is guaranteed to be generated by the probability distribution function $p(x)$.

Often than not, we want Gaussian random numbers such as $p(x) \propto te^{-x^2}$, $c(x) = \int_{0}^{x} e^{-x'^2} \, dx' = \text{erf}(x)$. There is an ad hoc trick for this kind of problem.

Box-Muller transform

Step 1: generate $u$ and $v$ uniformly in $(0, 1)$

Step 2: $x = \sqrt{-2\alpha \log u \cos(2\pi v)}, y = \sqrt{-2\alpha \log u \sin(2\pi v)}$
Claim: \(x, y\) are 2 independent Gaussian random numbers
Proof: Starting from a trivial equation

\[
P_{UV}(u,v)\,du\,dv = P_{XY}(x,y)\,dx\,dy
\]

\[
\implies \frac{\partial(u,v)}{\partial(x,y)} = \left( \frac{\partial(x,y)}{\partial(u,v)} \right)^{-1} = 1/\det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}
\]

\[
= 1/\det \begin{pmatrix} \frac{1}{2\alpha} \sqrt{-2\alpha/\log u \cos 2\pi v} & \frac{1}{2\alpha} \sqrt{-2\alpha/\log u \sin 2\pi v} \\ \sqrt{-2\alpha \log u (2\pi \sin 2\pi v)} & \sqrt{-2\alpha \log u (2\pi \cos 2\pi v)} \end{pmatrix}
\]

\[
= \frac{1}{\sqrt{2\pi \alpha}} e^{-\frac{x^2}{2\alpha}} \cdot \frac{1}{\sqrt{2\pi \alpha}} e^{-\frac{y^2}{2\alpha}} = P_X(x) \cdot P_Y(y)
\]

End of proof.

**Basics in Statistical Mechanics**

Let’s denote one general system state by \(\{\vec{r}_i, \vec{p}_i\} = \Gamma_i\)

\[
E(\{\vec{r}_i, \vec{p}_i\}) = E(\Gamma_i) = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i<j} V(||\vec{r}_i - \vec{r}_j||)
\]

For any thermodynamics variable \(f\), we have macro average value:

\[
< f(\{\vec{r}_i, \vec{p}_i\}) > = \frac{1}{Q_N} \int \prod_{i=1}^N d^3 p_i d^3 r_i f(\{\vec{r}_i, \vec{p}_i\}) e^{-\beta E(\{\vec{r}_i, \vec{p}_i\})}
\]

Or more concisely:

\[
< f(\Gamma) > = \frac{1}{Q_N} \int d\Gamma f(\Gamma) \rho(\Gamma) \quad \rho(\Gamma) = \rho^{eq}(\Gamma) = \frac{e^{-\beta E(\Gamma)}}{\int d\Gamma e^{-\beta E(\Gamma)}}
\]

in which, we denote the equilibrium distribution as \(\rho^{eq}(\Gamma)\)

**Metropolis Method**

First, let’s just consider Simple MC, (1) pick a series of \(\Gamma_i\) at random, (2) get ensemble average \(< f(\Gamma) > = \frac{1}{N} \sum_{i=1}^N f(\Gamma_i)\rho(\Gamma_i)\). This is officially correct for \(N \to \infty\), while highly efficient in practical. Then, we would turn to Importance Sampling MC: pick \(\Gamma_i\) with pdf close to \(\rho(\Gamma)\), and do biased sampling average. However it is still hard to do. Here, we introduce
the Metropolis Method. We would generate a sequence of states $\Gamma_i$ ($\Gamma_1 \rightarrow \Gamma_2 \rightarrow \Gamma_3 \rightarrow \cdots$), which, of course, are obtained by a certain rules. These sequent of states will always be called Markov Chain. We hope this sequence converges to physical states with pdf $\rho^{eq}(\Gamma)$. Then we can compute statical average to get the thermodynamics variable we interest:

$$< f(\Gamma) >= \frac{1}{N} \sum_{i=1}^{N} f(\Gamma_i)$$

To implement the Metropolis method in practice, we need to start form one configuration state($\Gamma_1$), which is deemed as a vector or a column of numbers $\rho_i = $ (probability that a member of the statistical ensemble of states is in some “i” state). Then the next job is to determine the transitions probability from state $i$ to state $j$: $\pi_{i \rightarrow j}$, which could be viewed as matrix element. We first observe 4 basic properties of this transition matrix: (1) Positivity, $\pi_{i \rightarrow j} \geq 0$, (2) Unity, $\sum_i \pi_{i \rightarrow j} = 1$, (3) Markov properties, $\pi_{i \rightarrow j}$ depende on $i$, not on earlier history, (4) Irreducibility or Ergodicty, for any two states $i, j$, there is at least one set of $\pi_{kk}$ such that, $\pi_{i \rightarrow k_1} \pi_{k_1 \rightarrow k_2} \cdots \pi_{k_n \rightarrow j} \neq 0$. The last critical property ensure us that “time average == ensemble average”.

One fundamental requirement is that, as said, we want the Markov process to converge to physical equilibrium. Or they do not destroy such an equilibrium distribution once it is reached. We call it Global Balance. Mathematically,

$$\rho^{eq}_i = \sum_j \rho^{eq}_j \pi_{j \rightarrow i}$$

The above equation tell us that matrix $\pi$ has a left eigenvector with eigenvalue 1, at least.

We could use a more finner version(in equilibrium) of balance to ensure the global balance, which is called Detailed Balance satisfying : (rate of transition $i \rightarrow j$) = (rate of transition $j \rightarrow i$). Or mathematically,

$$\rho^{eq}_i \pi_{i \rightarrow j} = \rho^{eq}_j \pi_{j \rightarrow i}$$

Easily, we could show that detailed balance (DB) $\implies$ global balance (GB):

$$\sum_j \rho^{eq}_j \pi_{j \rightarrow i} = \sum_j \rho^{eq}_i \pi_{i \rightarrow j} = \rho^{eq}_i \sum_j \pi_{i \rightarrow j} = \rho^{eq}_i$$

Let’s find the transition matrix $\pi_{i \rightarrow j}$, which satisfying detailed balance. Actually, We are entitled to do this quite freely since we only need the final equilibrium state regardless the the artificial un-physics process . Assuming that

$$\pi_{i \rightarrow j} = \alpha_{ij} \gamma_{i \rightarrow j}$$

in which, $\alpha_{ij}$ is the probability of selecting state $j$ if currently in state $i$, While $\gamma_{i \rightarrow j}$ is the probability of making the transition $i \rightarrow j$. Based on the same reason, we have the freedom
to choose $\alpha_{ij} = \alpha_{ji}$, although it might not true in physics. We then arrive
\[
\frac{\pi_{i \rightarrow j}}{\pi_{j \rightarrow i}} = \frac{\gamma_{i \rightarrow j}}{\gamma_{j \rightarrow i}} = \frac{\rho_j^eq}{\rho_i^eq} = e^{-\beta(E_j - E_i)}
\]
The last equal sign follows the canonical ensemble. We still have the luxury to choose our solution as long as it respect the Detailed Balance. One well-known solution is
\[
\gamma_{i \rightarrow j} = \begin{cases} 
1 & \text{if } E_i > E_j \\
\frac{\rho_j}{\rho_i} e^{-\beta(E_j - E_i)} & \text{if } E_i \leq E_j
\end{cases}
\]
The above solution is called **Metropolis Rule** (MR). Another possible solution we could use is
\[
\gamma_{i \rightarrow j} = \frac{\rho_i}{\rho_i + \rho_j}
\]
Here is a intuitive proof of MR: suppose we have an ensemble of $N_0$ systems and two states $E_i < E_j$, the number of transition $(j \rightarrow i) = N_{i \rightarrow j}$, the number of transition $(i \rightarrow j) = N_{j \rightarrow i}$. Following MR, we have
\[
N_{i \rightarrow j} = N_0 \rho_i \alpha_{ij} e^{-\beta(E_j - E_i)} \\
N_{j \rightarrow i} = N_0 \rho_j \alpha_{ij} = N_0 \rho_j \alpha_{ij}
\]
Therefor the net transition between these two states is
\[
\Delta N = N_{i \rightarrow j} - N_{j \rightarrow i} = N_0 \rho_i \alpha_{ij} \left( \frac{e^{-\beta E_j}}{e^{-\beta E_i}} - \frac{\rho_j}{\rho_i} \right) = N_0 \rho_i \alpha_{ij} \left( \frac{\rho_j^eq}{\rho_i^eq} - \frac{\rho_j}{\rho_i} \right)
\]
We draw the conclusion hter that if $\frac{\rho_j^eq}{\rho_i^eq} > \frac{\rho_j}{\rho_i}$, there will be a net transition from state $i \rightarrow j$, and vice-versa, This instructs the system towards the right direction. However, it does not promise that the system will converge to the equilibrium state, it might just oscillate around it, and then deviation could even be bigger and bigger. A more rigorous proof will be given on next lecture.

**Lecture 6**

**Review of Markov Chain**

Markov Chain: State 1 $\rightarrow$ 2 $\rightarrow$ 3 $\rightarrow$ $\cdots$, with probability of transition matrix,
\[
\pi_{i \rightarrow j} = \alpha_{ij} \gamma_{i \rightarrow j}
\]
We post constraint **Detail Balance** (DB) on this transition matrix, regardless it is not physically true process.

\[ \rho_{eq}^{i \rightarrow j} \pi_{j \rightarrow i} = \rho_{eq}^{j \rightarrow i} \pi_{i \rightarrow j} \]

In practice, we choose one convenient algorithm **Metropolis Rule** (MR) to fulfill the DB:

\[ \gamma_{i \rightarrow j} = \min \left( 1, \frac{\alpha_{ji} \rho_{eq}^{i \rightarrow j}}{\alpha_{ij} \rho_{eq}^{j \rightarrow i}} \right) = \min \left( 1, \xi \right) \]

Then we can check it really satisfying the DB,

\[ \frac{\pi_{i \rightarrow j}}{\pi_{j \rightarrow i}} = \frac{\alpha_{ij} \min(1, \xi)}{\alpha_{ji} \min(1, \xi^{-1})} = \left\{ \begin{array}{ll} \frac{\alpha_{ji}}{\alpha_{ij}} \frac{\xi}{1} & (\xi < 1) \\ \frac{\alpha_{ji}}{\alpha_{ij}} \frac{1}{\xi^{-1}} & (\xi \geq 1) \end{array} \right. = \frac{\alpha_{ji} \xi}{\alpha_{ij} \rho_{i}^{eq}} = \frac{\alpha_{ji} \xi}{\rho_{i}^{eq}} \]

So we conclude that the MR fulfill the DB, which respects GB. At the end of last lecture, we used a intuitive proof to show that DB Markov Chain will lead to equilibrium, which fail to be a real proof as it does not promise the chain will converge to the equilibrium. Now something glory, here is a more rigorous proof. (But still not be concrete enough for mathematician.)

**Convergence of Detailed Balance to Equilibrium**

It is is always good to understand the mechanism behind algorithm. Why Markow chain always lead to the equilibrium state? To answer it, we will borrow the idea used in Boltzman’s H-theorem. We start from a constructed variable \( G \), aimed to be a measurement of the deviation from equilibrium:

\[ G = \sum_{i} \frac{1}{\rho_{eq}^{i}} (\rho_{i} - \rho_{eq}^{i})^2 \]

From the above definition, we know clearly that \( G = 0 \) when the system is in equilibrium. Otherwise, \( G > 0 \). Rewrite \( G \) as,

\[ G = \sum_{i} \left( \frac{\rho_{i}^2}{\rho_{eq}^{i}} - 2 \rho_{i} + \rho_{eq}^{i} \right) = \sum_{i} \frac{\rho_{i}^2}{\rho_{eq}^{i}} - 1 \]

The last equal sign follows because \( \sum \rho_{i} = \sum \rho_{eq}^{i} = 1 \). Under general transition, as we know

\[ \rho_{i} \rightarrow \rho'_{i} = \sum_{m} \rho_{m} \pi_{m \rightarrow i} \]

\[ G \rightarrow G' = \sum_{mn} \pi_{m \rightarrow i} \pi_{n \rightarrow i} \frac{\rho_{m} \rho_{n}}{\rho_{eq}^{i}} - 1 \]
Detailed balance says that the transition matrix must follow the constraints

\[
\begin{align*}
\pi_{m \rightarrow i} &= \pi_{i \rightarrow m} \frac{\rho_i^{eq}}{\rho_m^{eq}} \\
\pi_{n \rightarrow i} &= \pi_{i \rightarrow n} \frac{\rho_i^{eq}}{\rho_n^{eq}}
\end{align*}
\]

\[\implies G' = \sum_{imn} \pi_{i \rightarrow m} \pi_{i \rightarrow n} \rho_i^{eq} \frac{\rho_m \rho_n}{\rho_m^{eq} \rho_n^{eq}} - 1\]

As for $G$, utilizing unity of transition matrix $\sum_m \pi_{i \rightarrow m} = \sum_n \pi_{i \rightarrow n} = 1$, we get

\[G = \sum_{im} \pi_{i \rightarrow m} \frac{\rho_i^2}{\rho_i^{eq}} - 1 = \sum_{mi} \pi_{m \rightarrow i} \frac{\rho_m^2}{\rho_m^{eq}} - 1\]

\[= \sum_{im} \pi_{i \rightarrow m} \frac{\rho_i^{eq}}{\rho_m^{eq}} \frac{\rho_m^2}{\rho_m^{eq}} - 1 = \sum_{im} \pi_{i \rightarrow m} \frac{\rho_i^{eq}}{\rho_m^{eq}} \rho_m^2 \sum_n \pi_{i \rightarrow n} - 1\]

\[= \sum_{imn} \pi_{i \rightarrow m} \pi_{i \rightarrow n} \rho_i^{eq} \left( \frac{\rho_m}{\rho_m^{eq}} \right)^2 - 1\]

In the second equal sign, we switch $i \leftrightarrow m$ since they are dummy index. And then, after applying DB in the third equal sign, we get the difference of $G$ before and after one step of transition:

\[\Delta G = G' - G = \sum_{imn} \pi_{i \rightarrow m} \pi_{i \rightarrow n} \rho_i^{eq} \left( \frac{\rho_m \rho_n}{\rho_m^{eq} \rho_n^{eq}} - \frac{\rho_m \rho_n}{\rho_m^{eq} \rho_n^{eq}} \right)\]

\[= \sum_{imn} \pi_{i \rightarrow m} \pi_{i \rightarrow n} \rho_i^{eq} \left( \frac{\rho_m \rho_n}{\rho_m^{eq} \rho_n^{eq}} - \frac{\rho_n \rho_m}{\rho_n^{eq} \rho_m^{eq}} \right)\]

Again we switch $n \leftrightarrow m$ since they are dummy index. Let’s add up the above two expression for $\Delta G$, and divide it by 2 so as to get back $\Delta G$,

\[\Delta G = -\frac{1}{2} \sum_{imn} \pi_{i \rightarrow m} \pi_{i \rightarrow n} \rho_i^{eq} \left( \frac{\rho_m}{\rho_m^{eq}} - \frac{\rho_n}{\rho_n^{eq}} \right)^2\]

We find that $\Delta G$ is negative definitive unless $\rho_m/\rho_m^{eq} = \rho_n/\rho_n^{eq}$ for $\pi_{i \rightarrow m}, \pi_{i \rightarrow n} \neq 0$. If for any $m$, $\pi_{i \rightarrow m} \neq 0$, then we would have:

\[\frac{\rho_m}{\rho_m^{eq}} = \text{const} = k \quad (\forall m)\]

or

\[\rho_m = k \rho^{eq}_m (\forall m)\]

\[\implies 1 = \sum_m \rho_m = k \sum_m \rho^{eq}_m = k\]
So we conclude that if all the states are directly connected to the current state (this is the ideal situation), $\Delta G = 0$ informs us that the system is in equilibrium, i.e.: $\rho_m = \rho_{eq}^m \forall m$. Otherwise, $G$ monotonically decreases. This illustrate that the system will always be approaching the equilibrium under DB transition.

More practically, $\pi_{i\rightarrow m} = 0$ happens for some states $m$, or say that not all the states are directly connected to current state, we may still have accidentally $\Delta G = 0 (G \neq 0)$. In this case, we only have $\rho_m = k\rho_{eq}^m (k \neq 1, \pi_{i\rightarrow m} \neq 0)$, which is obviously non-equilibrium state. Luckily, we know it must decrease later, since our system is defined to be ergodic. It must connected to any state within finite MC steps. So sooner or latter, when our simulation state explores more phase space, it will be in touch with that the states, then $G$ will again monotonically decrease towards zero, which means the system will keep approaching the equilibrium. We could also say ergodicy serves as a requirement here to ensure $G$ decreasing when the state is out of equilibrium.

Lecture 7

Monte Carlo Simulation for realistic Systems

We will cover the following issues in the future lectures:

1. What are the moves
2. Temperature variation
3. Periodicity
4. equilibrium
5. connection to thermodynamics
6. error analysis
7. molecules
8. other ensembles
9. free energy
1 MC Moves

pick an atoms i, make random displacement

\[ \Delta x_i = \Delta_0 (2 \times \text{ran()} - 1) \]
\[ \Delta y_i = \Delta_0 (2 \times \text{ran()} - 1) \]
\[ \Delta z_i = \Delta_0 (2 \times \text{ran()} - 1) \]

Two possible problems:

1. big moves usually rejected
2. small moves don’t do any thing

We want to choose such a displacement \( \Delta_0 \) that the accepted ratio is about 50%. Starting from a very small displacement, the calculated accepted ration will probably be \( \sim 0.99 \rightarrow \) increase \( \Delta_0 \rightarrow \) accepted ratio smaller \( \rightarrow \) keep increasing \( \Delta_0 \) until accepted ratio \( \sim \frac{1}{2} \)

In a more complicated situation, e.g.: working with dipole moment in spherical coordinate, we might have

\[ \vec{P} = \vec{P}(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \]

The MC moves could be

\[ \phi \rightarrow \phi + \Delta \phi \]
\[ \theta \rightarrow \cos \theta + \Delta (\cos \theta) \]

We choose to vary \( \cos \theta \) rather than \( \theta \), because the purpose here is to pick a uniform randomly point on a sphere. The area element in spherical coordinate is

\[ d\Omega = \sin \theta d\theta d\phi = d(\cos \theta) d\phi \]

One thing to point out is: when we have several coordinates to vary in one MC move, the order of them will not matter as long as we explore the whole phase space, which means that after long enough time running, they will alway be the same equilibrium.

2 Influence of Temperature

Generally:

- if \( k_b T > \text{typical } \Delta E \rightarrow \) easy to change system; most moves accepted
- if \( k_b T \ll \text{typical } \Delta E \rightarrow e^{-\Delta E/k_b T} \ll 1; \) most moves rejected
So the temperature is very key to the simulation. Start with a random configuration if cooling too quickly, it will go to nearest potential minimum state. which might not be desired the global minimum. The idea to solve this is called **Simulated Annealing**, which is an imaginary process invented to make a system to equilibrium state required. Let’s start our simulation from a high temperature, and decrease it step by step. This gives the system freedom to explore more phase space, or increase the probability of finding the global minimum.

Actually, the idea of annealing, actually artificial procedure, could be applied to any system with an analogy energy waiting to be minimized. For example:

1. Traveling salesman problem
2. Circuit network

### 3 Periodicity

The boundary could be any of these: rigid (reflecting), repulsive ($f \sim r^{-13}$) or periodic. Why bother? Because we want the bulk behavior rather than interface behavior. i.e.: the fraction of surface region

$$\frac{V_{\text{boundary}}}{V_{\text{cube}}} = \frac{6L^2 \rho}{L^3} \sim \frac{\rho}{L}$$

in which $L$ is the size of bulk, and $\rho$ is the width of boundary. It doesn’t matter in macroscopic level. While in microscopic level, $L \sim 10\rho$. In other words, 10% of the atoms are around surface, contrary to that we hope most of atoms show bulk properties.

---

**Lecture 8**

To remove the boundaries effect, we will introduce the periodic boundaries. Actually, we have two issues to take care:

1. atoms leaves the primary box, comes back from opposite side.
2. every atoms in primary box interact with nearest images of all others.

For first issue, we could code like these in Fortran ($box =$ size of primary box):

```fortran
if(xi.gt.box) xi = xi - box
if(xi.lt.box) xi = xi + box
```

For second issue, we could utilize that $0.5 \, box = hbox$ discriminate the two situation:
if(dx.gt.hbox) xi = dx - hbox
if(dx.lt.hbox) dx = dx + hbox

The above way of working periodicity is called “minimum image convention”, which requiring the cut-off length $R_c < hbox$.

One thing to notice is that “if” statement takes a lot the cpu time since the conditional statement is less efficient that the algorithm statement. To avoid these “if”s, we could employ following codes for first issue:

```plaintext
boxi2 = 2. / box
ix = x(i) * boxi2 - 1.0
x(i) = x(i) - ix * box
```

For second issue:

```plaintext
dx = x(2) - x(1)
idx = dx * boxi2
dx = dx - idx * box
```

In the above sample code, we notice that

$$idx = \begin{cases} 0 & |dx| < hbox \\ 1 & dx < hbox \\ -1 & dx < -hbox \end{cases}$$

4 Equilibration

Initial configurations are usually out of equilibration. We want to run it until it is in equilibration, by testing for stability of macroscopic variable to verify the equilibration.

5 Connecting to Statistical Mechanics

In NVT ensemble (fixed constant number of particles, volume and temperature of system)
The possible variables to calculate

- $E$: internal energy
- $c_v$: specific heat
- $s$: entropy
- $F$: Helmholtz free energy
- $p$: pressure
Starting from partition function

\[ Q_N = \frac{1}{N!} \int \prod_{i=1}^{N} \frac{d^3r_i d^3p_i}{h^3} e^{-\beta E} = \left( \frac{V^N}{N! \Lambda_T^{3N}} \right) \int \prod_{i=1}^{N} \frac{d^3r_i}{V} e^{-\beta U} \]

The first part is the classical ideal gas partition function, in which \( \Lambda_T = \frac{\hbar}{2\pi m kT} \) is called thermal wave length. The second part is the configuration partition due to the potential between each pair of particles, which is denoted \( Z_N \) below.

(i) Internal Energy \( E \)

\[
\langle E \rangle = \left( \sum_i \frac{p_i^2}{2m} + U_{LJ} \right) = \frac{3}{2} NKT + \left( \sum_{i<j} U_{ij} (|\vec{r}_i - \vec{r}_j|) \right) = -\frac{\partial}{\partial \beta} \log Q_N
\]

\[
\langle U_{LJ} \rangle = \frac{1}{N!} \int \prod_{i=1}^{N} \frac{d^3r_i d^3p_i}{h^3} U_{LJ} e^{-\beta E} Q_N = \frac{\int \prod_{i=1}^{N} d^3r_i U_{LJ} e^{-\beta U_{LJ}}}{\int \prod_{i=1}^{N} d^3r_i e^{-\beta U_{LJ}}} = -\frac{\partial}{\partial \beta} \log Z_N
\]

(ii) Specific heat \( C_v \)

\[
C_v = \left( \frac{\partial Q}{\partial \beta} \right)_V = \left( \frac{\partial E}{\partial \beta} \right)_T = \frac{3}{2} N K + \frac{\partial}{\partial T} \left( -\frac{\partial}{\partial \beta} \log Z_N \right)
\]

in which

\[
\frac{\partial}{\partial T} \left( -\frac{\partial}{\partial \beta} \log Z_N \right) = \frac{1}{kT^2} \frac{\partial^2}{\partial \beta^2} (\log Z_N) = \frac{1}{kT^2} \frac{\partial}{\partial \beta} \left[ \frac{\int U e^{-\beta U}}{\int e^{-\beta U}} \right]
\]

\[
= \frac{1}{kT^2} \left[ \frac{\int U^2 e^{-\beta U}}{\int e^{-\beta U}} - \left( \frac{\int U e^{-\beta U}}{\int e^{-\beta U}} \right)^2 \right]
\]

\[
= \frac{1}{kT^2} (\langle U^2 \rangle - \langle U \rangle^2) = \frac{1}{kT^2} (\Delta U)^2
\]

So finally, the specific heat is express by

\[
C_v = \frac{3}{2} N K + \frac{1}{kT^2} (\Delta U)^2
\]

Lecture 9

(iii) Entropy \( S \)

We sometimes want the entropy of the system, especially the difference of entropy between two states. One standard method is through thermodynamics integration. For a reversible
process with fixed volume

\[ dS = \frac{\delta Q}{T} = \frac{dE}{T} = \frac{c_v}{T} dT \]

\( \Rightarrow \)

\[ \Delta S = \int_1^2 dS = \int_1^2 \frac{c_v(T)}{T} dT \]

Generally, \( c_v(T) \) is a function of temperature. Within MC, we express it as

\[ \Delta S = \sum_i \frac{c_v(T_i) \Delta T}{T_i} \]

(iv) **Helmholtz free energy** \( F \)

Based on above entropy difference, we are almost to get Helmholtz free energy:

\[ \Delta F = \Delta(U - TS) = \Delta U - (T_2 S_2 - T_1 S_1) = \Delta U - T_2 \Delta S + \Delta T S_1 \]

So all the variables appearing above could be calculable except \( S_1 \). So only when knowing \( S_1 \), which is just the entropy of the beginning state, we would be able to get desired \( F \). For convenience, state 1 could be chosen as ideal gas state (at large temperature), whose entropy is known analytically.

We will introduce several more detailed methods to compute free energy in lecture 12.

(v) **Pressure** \( p \)

We introduce an important theorem **Virial Theorem**. For a system of \( N \) particles, using canonical ensemble,

\[ H (\{ \vec{r}_i, \vec{p}_i \}) = \sum_{i=1}^N \frac{p_i^2}{2m} + U(|\vec{r}_i - \vec{r}_j|) \]

First, we point out that from classical mechanics, if \( U \) is independent of time, the the above hamiltonian is just identical with energy i.e.: \( H = E \), as in most cases. The probability of a specific configuration \( \{ \vec{r}_i, \vec{p}_i \} \) is

\[ \text{Prob} (\{ \vec{r}_i, \vec{p}_i \}) = \frac{1}{Q} e^{-\beta H} d\Gamma \]

in which, \( Q \) is partition function, \( d\Gamma \) is the phase space element. Explicitly,

\[ Q = \int e^{-\beta H} d\Gamma \]

\[ d\Gamma = \frac{1}{N!} \prod_{i=1}^N \frac{d^3 r_i d^3 p_i}{h^3} \]
For any two variables \( x_i, x_j \) from coordinates or momentum \( \{ \vec{r}_i, \vec{p}_i \} \), we have the ensemble average

\[
\left\langle x_i \cdot \frac{\partial H}{\partial x_j} \right\rangle = \int d\Gamma \frac{1}{Q} x_i \frac{\partial H}{\partial x_j} e^{-\beta H}
= -\frac{1}{\beta} \int d\Gamma \frac{1}{Q} x_i \frac{\partial}{\partial x_j} e^{-\beta H}
= -\frac{1}{\beta} \int \frac{1}{Q} \left[ \frac{\partial}{\partial x_j} (x_i e^{-\beta H}) - \delta_{ij} e^{-\beta H} \right]
\]

After integration, the first term in square bracket is just boundary term, and will be zero since: Case I, \( x_i \) = some momentum, then \( e^{\beta \alpha H} \to 0 \) (\( p_i \to \infty \)) Case II, \( x_i \) = some coordinate, if the system is confined in a certain space, then \( e^{-\beta H} \to 0 \) (\( r_i \to \infty \)); if the system has a periodic boundary condition, then the two limits just cancels. So what’s left over will be

\[
\left\langle x_i \cdot \frac{\partial H}{\partial x_j} \right\rangle = kT \delta_{ij} \int \frac{d\Gamma}{Q} e^{-\beta H} = kT \delta_{ij}
\]

The above equation is the generalized form of Equal Partition Theorem.

**Example**

(1) \( x_i = x_j = p_{i\alpha} \), \( H = \frac{p_{i\alpha}^2}{2m} + U \)

\[
\frac{\partial H}{\partial x_j} = \frac{p_{i\alpha}}{m} \implies \left\langle x_i \cdot \frac{\partial H}{\partial x_j} \right\rangle = \left\langle \frac{p_{i\alpha}^2}{m} \right\rangle = kT
\]

or

\[
\left\langle \frac{1}{2} m_i v_{i\alpha}^2 \right\rangle = \frac{1}{2} kT
\]

(2) \( x_i = x_j = r_{i\alpha} \), \( H = \frac{p_{i\alpha}^2}{2m} + U \)

\[
\frac{\partial H}{\partial x_j} = \frac{\partial H}{\partial r_j} = \frac{\partial U}{\partial r_j} = -F_{i\alpha}
\]

or

\[
\left\langle \vec{r}_{i\alpha} \vec{F}_{i\alpha} \right\rangle = -kT
\]

Sum over all particles

\[
\left\langle \sum_i \vec{r}_i \cdot \vec{F}_i \right\rangle = -3NkT
\]

We need to write the force term into two parts: inter molecular force and the force from boundaries

\[
\vec{F}_i = \vec{F}_i^{int} + \vec{F}_i^{ext}
\]
in which the inter molecular force come from the Lennard-Jones potential between each pair of particles.

\[ \vec{F}_{int}^i = -\frac{\partial}{\partial r_i^i} \sum_{j \neq i} V_{LJ}(|\vec{r}_i^i - \vec{r}_j^j|) \]

\[ = -\sum_{j \neq i} V'_{LJ}(r_{ij}) \frac{\vec{r}_i^i - \vec{r}_j^j}{|\vec{r}_i^i - \vec{r}_j^j|} \]

\[ = \sum_{j \neq i} \vec{F}_{ij} \]

So the first part is

\[ \left\langle \sum_i \vec{r}_i \cdot \vec{F}_{int}^i \right\rangle = \left\langle \sum_i \vec{r}_i \cdot \sum_{j \neq i} \vec{F}_{ij} \right\rangle \]

\[ = \sum_{ij} ' \left\langle \vec{r}_i \cdot \vec{F}_{ij} \right\rangle = \frac{1}{2} \sum_{ij} ' \left\langle \vec{r}_i \cdot \vec{F}_{ij} + \vec{r}_j \cdot \vec{F}_{ji} \right\rangle \]

\[ = \frac{1}{2} \sum_{ij} ' \left\langle (\vec{r}_i - \vec{r}_j) \cdot \vec{F}_{ji} \right\rangle = \frac{1}{2} \sum_{ij} ' \left\langle \vec{r}_{ij} \cdot \vec{F}_{ij} \right\rangle \]

\[ = - \sum_{i<j} \left\langle \vec{r}_{ij} \cdot U'(r_{ij}) \right\rangle \]

The second part will be

\[ \left\langle \sum_i \vec{r}_i \cdot \vec{F}_{ext}^i \right\rangle = \int_{\partial V} \vec{r}(-p) d\vec{s} = (-p) \int_V \nabla \cdot \vec{r} dV = -3pV \]

To sum up

\[ \therefore \sum_{i<j} \left\langle \vec{r}_{ij} \cdot U'(r_{ij}) \right\rangle - 3pV = -3NkT \]

or

\[ p = \rho kT - \frac{1}{3N} \sum_{i<j} \left\langle \vec{r}_{ij} \cdot U'(r_{ij}) \right\rangle \]

**Lecture 10**

2 kind molecules bonds:

- Soft bonds \( \Delta |\vec{r}_i - \vec{r}_j| \sim \sigma \) or variable length

- Stiff bonds \( \Delta |\vec{r}_i - \vec{r}_j| \ll \sigma \) or fixed length
For soft bonds, we adopt the following form of potential for the whole system.

\[ V(\{\vec{r}_i\}) = \sum_{i<j} V_{LJ}(\vec{r}_i - \vec{r}_j) + \sum_{\text{bonds}} V_b(\vec{r}_i - \vec{r}_j) \]

The bonds potential between two atoms could be either harmonic or FENE

\[ V_b(r_{ij}) = \frac{1}{2} k (r_{ij} - r_b)^2 \quad (\text{Hooke's Law}) \]

Or

\[ V_b(r_{ij}) = -\frac{1}{2r_b^2} \log \left( 1 - \frac{r_{ij}^2}{r_b^2} \right) \quad (\text{FENE}) \]

There will be 2 kinds of MC moves (1) molecule moves; (2) bonds moves.

**Pseudo-code (l = length of molecule)**

```
do i=1, ncycle
  do j=1, nmol
    pick a random molecule
    r = random number
    m = (l+1) * r
    if( m==0 ) move molecule
    else move atom m
    U differenced calculated
    test via MC method
  end do
end do
```

For hard bonds, we view the whole molecule as a unit, which could only do translation and rotation movement. For a simple case, i.e. dima rotation. Let’s make a unit vector \( \hat{u} \propto \vec{r}_i - \vec{r}_j \). We would like to randomly rotate this unit vector. First is to generate another random vector \( \hat{v} \), then we could get the new \( \hat{u} \):

\[ \hat{u} = \frac{\hat{u} + \gamma \hat{v}}{\sqrt{1 + 2\gamma \hat{u} \cdot \hat{v} + \gamma^2}} \]

So the question now is how to generate a random vector \( \hat{v} \)

**Method I**

random unit vector \( \hat{v} \rightarrow \) random \( \theta \) & \( \phi \). We observe that the surface element in spherical coordinate is \( r^2 \sin \theta \ d\theta \ d\phi = d(\cos \theta) d\phi \). So for true uniform random orientation
distribution, we need to call two random number \( r_1, r_2 \in (0, 1) \)

\[
\begin{align*}
\phi &= 2\pi r_1 \\
\cos \theta &= 2r_2 - 1
\end{align*}
\]

**Method II**

**Method III**

**Lecture 11**

Canonical (or NVT) ensemble, although equivalent to other ensembles in thermal-dynamics limits, is not always optimal for every cases. Sometimes, we shall adopt other ensemble for convenience. Here we first recall what is so called NVT ensemble, and then introduce the NPT ensemble.

I. NVT Ensemble

First, let’s review the NVT ensemble, which could be started from the fundamental postulate. The fundamental postulate of statistical mechanics says that for an isolated system in equilibrium, all microstates with same energy are equally likely. This tells us that the macrostate corresponds to the state with maximal number of microstates. Consider a thermodynamic system interacting with a reservoir, which supplies energy. For NVT ensemble, which means \( N, V \) fixed and \( E \) vary, the probability to have a specific energy is

\[
\text{Prob}(E) \propto \text{number of microstates with energy } E = \Omega(E) \Omega'(E')
\]

in which the primed the variable is for the reservoir and the total energy is the summation of the interested system and reservoir, plus the interaction between them.

\[
E_{\text{tot}} = E(\{x\}) + E'(\{\vec{x}'\}) + \Delta E(\{\vec{x}, \vec{x}'\})
\]

in which

\[
\begin{align*}
E &\propto V \\
E' &\propto V' \\
\Delta E &\propto V^{2/3} \ll E \text{ or } E'
\end{align*}
\]

Thus the interaction could be ignored. We have entropy for our system:

\[
S = k_b \log \Omega(E)
\]
and for the reservoir is

\[ S'(E') = k_b \log \Omega'(E') = k_b \log \Omega'(E_{\text{tot}} - E) \]

\[ = S'(E_{\text{tot}}) - \frac{\partial S'}{\partial E'} E + \frac{1}{2} \frac{\partial^2 S'}{\partial E'^2} E^2 \]

\[ \approx S'(E_{\text{tot}}) - \beta E + \mathcal{O}(E^2/E') \]

in which, we set the reservoir's characteristic variable as

\[ \left( \frac{\partial S'}{\partial E'} \right)_{N,V} = \beta \]

And also to notice that first term above is a constant, the second term is \( \frac{\partial^2 S'}{\partial E'^2} E \sim \mathcal{O}(E) \) \( \mathcal{O}(E) \sim \mathcal{O}(E) \) \( \mathcal{O}(E) \) \( \mathcal{O}(E) \), while the last term is \( \frac{\partial^2 S'}{\partial E'^2} E^2 \sim \mathcal{O}(E^2) \mathcal{O}(E^2) = \mathcal{O}(E^2) \mathcal{O}(E^2) = \mathcal{O}(E) \). or say, we could neglect all high order terms. Finally, we get the desired result

\[ \text{Prob}(E) = \Omega(E) \Omega'(E_{\text{tot}} - E) \propto \Omega(E) e^{-\beta E} \]

And the normalized results is

\[ P_r(E) = \Omega(E) \frac{e^{-\beta E}}{Q_{\text{NVT}}} \]

in which \( Q_{\text{NVT}} \) is called the \( NVT \) partition function, defined to be the summation of all microstates' probability

\[ Q_{\text{NVT}} = \sum_r e^{-\beta E_r} = \int dE \Omega(E) e^{-\beta E} \]

II. \( NPT \) Ensemble

Secondly, for \( NPT \) ensemble, we have fixed \( N \) but \( E, V \) vary.

\[ \text{Prob}(E, V) \propto \text{number of microstates with energy } E \text{ and volume } V \]

\[ = \Omega(E, V) \Omega'(E', V') \]

And similarly,

\[ S'(E', V) = k_b \log \Omega'(E', V') = k_b \log \Omega'(E_{\text{tot}} - E, V_{\text{tot}} - V) \]

\[ \approx S'(E_{\text{tot}}) - \frac{\partial S'}{\partial E'} E - \frac{\partial S'}{\partial V'} V \]

\[ \approx S'(E_{\text{tot}}) - \beta E - \beta p V \]

in which, we also set the reservoir's characteristic variable as

\[ \left( \frac{\partial S'}{\partial E'} \right)_{N,V} = \beta \]

\[ \left( \frac{\partial S'}{\partial V'} \right)_{N,E} = \beta p \]
So for the \textit{NPT} ensemble the probability to be in a specific \((E,V)\) state is
\[
P_r(E,V) = \Omega(E,V) \frac{e^{-\beta(E+pV)}}{Q_{NPT}}
\]
in which \(Q_{NVT}\) is called the \textit{NPT} partition function, defined to be the summation of all microstates’ probability
\[
Q_{NPT} = \int dE \int dV \Omega(E,V) e^{-\beta(E+pV)}
\]

\textbf{Fulfillment in MC}

In MC, what we want is the transition probability matrix: \(\pi_{j \rightarrow i}\)
\[
\rho_i \rightarrow \rho_i' = \sum_j \rho_j \pi_{j \rightarrow i}
\]

\textbf{Lecture 12}

\textbf{Helmholtz Free Energy} \(F\) \textit{in NVT ensemble}

It is not possible to measure the the free energy directly in a simulation. We need to introduce several methods to compute \(F\) here. Firs let’s review the methods stated before.

\textbf{Review}

Simply, we want the difference of free energy between two states: \(F_1\) is of the interested state and \(F_2\) is of the reference state. So as learned before, \(\Delta F = F_1 - F_2\) could be get by

(a) \(\Delta F = \Delta U - T_2 \Delta S + \Delta TS_1\);

(b) \(F = N\mu\) in which \(\mu\) could be got by particle insertion method;

(c) \(\Delta F = -k_b T \log \frac{Q_1}{Q_0} = -k_b T \log \langle e^{-\beta \Delta U}\rangle\).

All of them above has this or that defects in practice. We try to extend them to better methods in the following sections.

\textbf{Thermodynamics Integration Method}

This method is the generalization of the method (a) above. The weak point of (a) is that we need to get entropy difference , which might not be easy.
\[
\Delta S = \int_{T_1}^{T_2} \frac{dU}{T} = \int_{T_1}^{T_2} \frac{C_v(T)dT}{T}
\]
Another thing is that we need a well known initial entropy, which could only be ideal gas state. To get ride of the weak point, we first notice that

\[ dF = -SdT - pdV \quad \implies \quad \begin{cases} \left( \frac{\partial F}{\partial T} \right)_v = -S \\ \left( \frac{\partial F}{\partial v} \right)_T = -p \end{cases} \]

Entropy could not be measured directly, which prevent us integrating directly. However,

\[ \begin{cases} \left( \frac{\partial F/T}{\partial v/T} \right)_v = U \\ \left( \frac{\partial F/T}{\partial v/T} \right)_v = -p/T \end{cases} \quad \implies \quad d\left( \frac{F}{T} \right) = U d\left( \frac{1}{T} \right) - \frac{p}{T} dV \]

the pressure and energy are mechanical quantities, they could be measured directly either in experiment or computer simulation. Then the connection between two state is:

\[ \frac{F_2}{T_2} - \frac{F_1}{T_1} = \int_1^2 \left[ U d\left( \frac{1}{T} \right) - \frac{p}{T} dV \right] \]

So to compute the free energy of a interested state, all we need is to connect it to a well known state (e.g.: ideal gas state whose free energy is analytically known) by a constructed reversible path on phase diagram. Then the change in \( F \) along the path could be simply calculated by the summation or integration.

Actually, the above path is not necessarily to be physically meaningful. Any artificially constructed path could be applied in the thermodynamics integration method. For example, let’s construct the potential energy with one parameter \( \lambda \):

\[ U(\lambda) = \lambda U_I + (1 - \lambda) U_{II} \]

Obviously, \( U(0) = U_{II} \) and \( U(1) = U_I \). And as we know

\[ Q(\lambda) = \frac{1}{N! \Lambda^{3N}} \int \prod_i^N d^3 r_i e^{-\beta U(\lambda)} \quad \text{and} \quad F(\lambda) = -k_B T \log Q(\lambda) \]

The derivative of \( F(\lambda) \) with respect to \( \lambda \) can be written as an ensemble average:

\[ \left( \frac{\partial F(\lambda)}{\partial \lambda} \right)_{NVT} = -\frac{1}{\beta} \frac{\partial}{\partial \lambda} \log Q(\lambda) = -\frac{1}{\beta Q} \left( \frac{\partial Q(\lambda)}{\partial \lambda} \right) \]

\[ = -\frac{1}{\beta Q N! \Lambda^{3N}} \int \prod_i^N d^3 r_i \left( -\beta \frac{\partial U}{\partial \lambda} \right) e^{-\beta U(\lambda)} \]

\[ = \frac{\int \prod_i^N d^3 r_i \left( -\beta \frac{\partial U}{\partial \lambda} \right) e^{-\beta U(\lambda)}}{\int \prod_i^N d^3 r_i e^{-\beta U(\lambda)}} \]

\[ = \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{U(\lambda)} \]
where \( \langle \cdots \rangle_{U(\lambda)} \) denotes an ensemble average for a specific potential energy \( U(\lambda) \). After that, we would be able to get the free energy difference between these two systems,

\[
F(1) - F(0) = \int_0^1 d\lambda \mathcal{A} \left( \frac{\partial U(\lambda)}{\partial \lambda} \right)_{U(\lambda)}
\]

One extra point to see is that, since \( \frac{\partial U(\lambda)}{\partial \lambda} = U_I - U_{II} \), we have the following definite conclusion:

\[
\frac{\partial^2 F}{\partial \lambda^2} = -\beta \left[ \langle (U_I - U_{II})^2 \rangle - \langle (U_I - U_{II}) \rangle^2 \right] \leq 0
\]

So the slop of function \( F(\lambda) \) must decrease with respect to \( \lambda \), which is just a convenient check point to see whether our code is right or not.

**Overlapping Distribution Method**

Supposing 2 systems in \( NVT \) ensemble with potentials \( U_1 \) and \( U_0 \), from method (c)

\[
\Delta F = -k_b T \log \frac{Q_1}{Q_0} = -k_b T \log \langle e^{-\beta \Delta U} \rangle_0
\]

in which \( \Delta U = U_1 - U_0 \), and \( \langle \cdots \rangle_0 \) denotes the ensemble average by system 0. However \( \Delta U \) could be very large or very small, while not necessarily following the distribution of \( NVT \) ensemble distribution, which make the MC calculation quite inefficient. To improve, we could simply divide the potential into several steps, e.g.:

\[
U_0 \rightarrow U_a \rightarrow U_b \rightarrow U_c \rightarrow \cdots \rightarrow U_1
\]

Then the result would be like

\[
\Delta F = -k_b T \log \prod_i \frac{Q_i}{Q_{i-1}} = -k_b T \log \prod_i \langle e^{-\beta \Delta U_i} \rangle_i
\]

If \( \beta \Delta U_i \ll 1 \implies \Delta F = -\sum_i \Delta U_i \) One more better method is called Overlapping Distribution Method. Define \( P_i(\Delta U) \) as pdf of \( \Delta U \) in system with energy \( U_1 = U_0 - \Delta U \)

**Umbrella Sampling Method**
References


