

# 6. Computer Simulation Methods

## ▼ Table of Contents

### Table of Contents

#### 6.1. Introduction

##### 6.1.1. Time Averages, Ensemble Averages and Some Historical Background.

##### 6.1.2. A Brief Description of the Molecular Dynamics Method

##### 6.1.3. The Basic Elements of the Monte Carlo Method

##### 6.1.4. Differences Between MD and MC

#### 6.2. Calculation of Thermodynamic Properties

##### 6.2.5. Radial Distribution Functions

#### 6.3. Phase Space

#### 6.4. Practical Aspects of Computer Simulation

##### 6.4.2. Choosing the Initial Configuration

#### 6.5. Boundaries

##### 6.5.1. Periodic Boundary Condition

##### 6.7 Truncating the Potential and the Minimum Image Convention

##### 6.7.3. Problems with Cutoffs and How to Avoid Them

#### 6.8. Long-range Forces

##### 6.8.1. The Ewald Summation Method

## 6.1. Introduction

- Energy minimisation generates individual minimum energy configurations of a systems
  - can predict accurately the properties of the system.
  - can derive partition function from which thermodynamic properties can be calculated.
  - mainly for small molecules.
- Computer simulation methods enable us :
  - to study systems in the condensed phase
  - predict their properties through use of techniques through small replications of macroscopic system with manageable # of atoms.
- A simulation generate representative configuration of small replication in a way that accurate values of structural and thermodynamic properties can be obtained.
  - enables time-dependent behavior of atomic and molecular systems
  - providing detail picture of how one conformation/configuration changes to another.

### 6.1.1. Time Averages, Ensemble Averages and Some Historical Background.

- To determine value of a property of system, pressure or heat capacity:
  - these properties depend on the position and momenta of N particles that comprise the system.
  - Instantaneous value of property A is written as  $A(p^N(t), r^N(t))$ , representing momenta and position at time t
    - Instantaneous value can fluctuate over time as a result of interactions between particles.
  - **Time average** — average of A over the time of measurement.

$$A_{ave} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(p^N(t), r^N(t)) dt \quad (6.1)$$

- To calculate average values of properties in the system, necessary to simulate the dynamic behavior.
  - For any rearrangement of atoms, to get the force acting on it from interactions with other atoms, one would differentiate the energy function.
  - from force, can determine the acceleration via Newton's 2nd law.
  - Integration of equations of motion would give trajectory that describes how the positions, velocities, and accelerations of particles vary with time.
- Limitation:
  - For macroscopic system ( $10^{23}$ ), difficult to get initial configuration and even the integrations of motion.
  - That's why stat mech exist — in which a single system evolve in time is replaced by large number of replications of the system that are considered simultaneous.
  - Average is replaced by *ensemble average*, or expectation value: average value of property A over all replications of the ensemble generated by the simulation

$$\langle A \rangle = \iint dp^N dr^N A(p^N, r^N) \rho(p^N, r^N) \quad (6.2)$$

- value of property A over all replications of the ensemble generated by simulation.
- last part is the probability density of the ensemble — probability of finding a configuration with momenta  $p^N$  and positions  $r^N$ .
- Under conditions of constant # of particles, volume and temperature, the probability is the familiar Boltzmann distribution.

-- --

$$\rho(p^N, r^N) = \frac{\exp(-\beta E(p^N, r^N))}{Q} \quad (6.3)$$

### 6.1.2. A Brief Description of the Molecular Dynamics Method

- Molecular Dynamics calculates the ‘real’ dynamics of the system, where time averages of properties can be calculated.
  - Atomic position can be derived by integration of Newton’s equations of motion.
  - a *deterministic* method — where the state of the system at the future can be presented from the present.

#### Hard sphere model

- Behavior of particles is similar to pool, one ball is moving to collide with another ball and once collided, new velocity is assigned. Not quite like atoms.
- The force does not change until collided.

#### Atomic model

- LJ potential, the force between 2 atoms or molecules changes **continuously** with their separation.
- Need to integrate in very short time step.



**Timestep** — the interval in which the Newton’s equations of motions are being integrated.

- At each timestep, the forces on each atoms are computed and combined with current positions and velocities to generate new position and velocities a short time ahead.
- Force within an interval is considered to be constant.

- Molecular dynamics simulation generates a *trajectory* that describes how the dynamic variables change with time.

$$\langle A \rangle = \frac{1}{M} \sum_{i=1}^M A(p^N, r^N) \quad (6.5)$$

- M is # of timesteps.

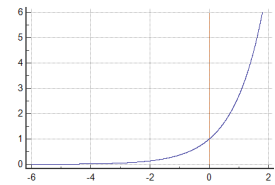
### 6.1.3. The Basic Elements of the Monte Carlo Method

- MD simulation successive configuration is connected through time. NOT MC.
- Monte Carlo simulation, each configuration depend upon its predecessor and not upon other configurations previous visits.
  - generates configurations randomly, and uses a special criteria to decide rather or not to accept each configuration.

- The probability of obtaining a given configuration is equal to its Boltzmann factor.  
 $\exp\{-V(r^N)/k_B T\}$
- average of these properties is obtained by averaging the # of values calculated, M.

$$\langle A \rangle = \frac{1}{M} \sum_{i=1}^M A(r^N) \quad (6.6)$$

- Metropolis MC is most popular
- MC, new config generated by randomly moving a single atom/molecule, several atoms/molecules, or rotating bonds,
  - Energy of new config is calculated using potential energy function.
    - If  $E(\text{new}) < E(\text{old})$ , accept config.
    - If  $E(\text{new}) > E(\text{old})$ , then Boltzmann factor of the energy difference is calculated
      - $\exp[-\beta(V_{\text{new}}(r^N) - V_{\text{old}}(r^N))]$
      - Another random #  $0 < x < 1$  is generated to compare with the Boltzmann factor
        - If random #  $>$  Boltzmann factor, move is rejected and original position retained
        - If random #  $<$  Boltzmann factor, accept new configuration
        - The purpose of this is to permit moves to states of higher energy.
        - $\downarrow$  uphill move  $\rightarrow \uparrow V_{\text{new}}(r^N) - V_{\text{old}}(r^N) \rightarrow \uparrow$  probability of getting accepted.



look at negative side

#### 6.1.4. Differences Between MD and MC

| MD  | MC  |
|---|---|
| time dependence of properties of the system   | no temporal relationship between successive configs               |
| configuration of the system at any time in the future can be predicted by any time in the past. | outcome of each trial depends only upon its immediate predecessor |
| Kinetic energy contribution to total energy   | total energy determined directly from potential energy function   |
| traditional in NVE; common in NPT, $\mu$ VT   | traditional in NVT  |

## 6.2. Calculation of Thermodynamic Properties

## 6.2.5. Radial Distribution Functions

- RDF is useful way to describe structure of a system, particularly of liquids.



RDF,  $g(r)$ , gives the probability of finding an atom (or molecule, if simulating a molecular fluid) a distance  $r$  from another atom (or molecule) compared to the ideal gas distributions.

- dimensionless
- 
- Radial distribution function of a liquid is intermediate between solid and gas
    - small # of peaks in short distances
    - steady decay to constant value at longer distances.
    - For short distances (less than atomic diameter)  $g(r) = 0$ 
      - due to strong repulsive forces.
    - first peak at  $g(r) \approx 3$ 
      - this means 3x likely that 2 molecules would have this distance separation.
    - where peak falls and passes through minimum value means the chances of finding two atoms with this separation are less.
    - At long distances,  $g(r)$  tends to the value, which means there is no long-range order.
  - rdf is kind of like a histogram.

## 6.3. Phase Space



**Phase space** is when all the possible states of a system is being explored. It consists of  $3N$  coordinates for both position and momenta

- For a system with  $N$  atoms,  $6N$  values are required to define the state of the system. (3 coordinates per position and 3 coordinates per momenta)
- MD generate sequence of points in phase space connected in time.
- MC only has  $3N$ -dimensional space corresponding to the positions of the atoms. No momenta
- If it is possible to visit all points in the phase space., then the partition function could be calculated by summing the values of  $\exp(-\beta E)$ .



**Ergodic** refers to a phase-space trajectory, in which all possible states of a system is explored.

- simulations cannot achieve ergodic trajectory because phase space is longer than the age of the universe, and then, can it provide the estimate of ‘true’ energy and other thermodynamic properties.

### Mechanical Properties

- mechanical properties* like internal energy, pressure, and heat capacity are routinely obtained in MD/MC.
- mechanical properties comes from derivative of the partition function.

$$U = \frac{k_B T^2}{Q} \frac{\partial Q}{\partial T} \quad (6.20)$$

- Let’s look at 6.20.  $Q$  is substituted with the NVT partition function, finally written as

$$U = \int \int dp^N dr^N \rho(p^N, r^N) \frac{A(p^N, r^N)}{Q} = \frac{1}{Q} \int \int dp^N dr^N \exp\left(-\frac{E(p^N, r^N)}{k_B T}\right) E(p^N, r^N) \rho(p^N, r^N) \quad (6.24)$$

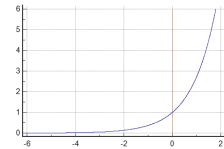
- High values of  $E(p^N, r^N)$  have very low probability and make an insignificant contribution to this integral.
  - MD MC preferentially generate states of low energy, which are the states that make a significant contribution to the integral in 6.24.
  - These method sample phase space in a way that is representative of the equilibrium states and can generate accurate estimate of mechanical properties.

### Thermal Properties

- Other thermodynamic properties, difficult to determine accurately without special techniques.
  - entropic and thermal properties:* free energy, chemical potential, and entropy.
- thermal properties relates to partition function itself.

$$A = -k_B T \ln Q \quad (6.21)$$

- Let’s look at Helmholtz free energy.



look at positive side



Configurations with a **high energy make a significant contribution** to the integral due to the presence of the exponential term  $\exp\left(-\frac{E(p^N, r^N)}{k_B T}\right)$ .

- Because MD MC preferentially samples the lower-energy regions of phase space.
- Ergodic trajectory would of course visit all these high-energy regions, which would not be adequately sample by a real simulation.

## 6.4. Practical Aspects of Computer Simulation

- Simulation usually performed with relatively large # of atoms over many iterations or time steps.

- intra- and intermolecular # of atoms and interaction described is empirical — molecular mechanics
- Faster computer and new theoretical techniques enable simulations to be performed by QM or mixed.
- First step is to decide which energy model is to be used to describe the interactions within the system.
- *equilibration phase* is performed, where system evolved from initial config until stability is achieved.

#### 6.4.2. Choosing the Initial Configuration

- Important to choose a configuration that is close to the state which it is desired to simulate.
- Initial configuration does not contain any high-energy interactions that would cause instabilities in the simulation. Just do minimization before hand.

### 6.5. Boundaries

- correct treatment of boundaries and boundary effects is crucial to simulation methods because it enables ‘macroscopic’ properties to be calculated from simulations using small # of particle.

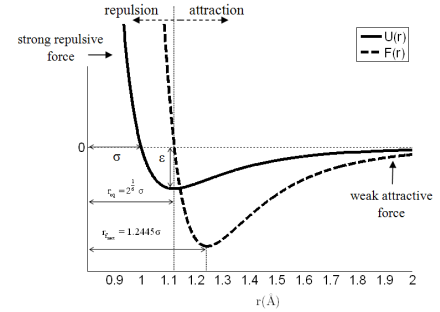
#### 6.5.1. Periodic Boundary Condition

- enable simulation to be performed in a small # of particles, like it is in bulk fluid.
  - it is replicated in all directions to give a periodic array.
  - In 2D, each box is surrounded by 8 neighbors
  - In 3D, each box is surrounded by 26 neighbors.
- Even if a particle leaves the box and enters into the neighboring box, the neighboring box can come back to the central box that the number remains constant.
- Important to simulate a cell close to the molecule’s geometry
  - for example: not rectangular prism, no spherical shapes.
  - truncated octahedron and rhombic dodecahedron provide periodic cells that approx spherical
- There for some simulation where it is inappropriate to use standard periodic boundary conditions in all directions.
  - studying adsorption of molecules onto a surface, inappropriate to use the usual periodic boundary conditions for motion perpendicular to the surface.
    - Rather, surface is modelled as a true boundary, for example by explicitly including atoms in the surface.
- Limitation:
  - not possible to achieve fluctuations that have wavelength greater than the length of the cell.

### 6.7 Truncating the Potential and the Minimum Image Convention

- The most time-consuming part of a MD MC simulation is the calculation of non-bonded energies and/or forces.

- number of bond-stretching, angle-bending, and torsional terms in a force field model are all proportional to the number of atoms
- number of non-bonded term that need to be evaluated increases in the order of  $N^2$ .
- Non-bonded interactions are calculated between every pair of atoms in the system.



- To deal with non-bonded interactions is to use a *non-bonded cutoff* and apply the *minimum image convention*.
- LJ potential falls off very rapidly with distance: at  $2.5\sigma$  LJ has 1% of its value at  $\sigma$ .
  - This reflects the  $r^{-6}$  distance dependence of the dispersion interaction.
- In minimum image convention, each atom ‘sees’ at most just one image of every other atom in the system (repeated infinitely via periodic boundary method)
- When cutoff is used, the interactions between all pairs of atoms that are further apart than the cutoff value are set to zero, taking into account the closest image.
- cutoff should not be so large that a particle sees its own image or the same molecule twice.

### 6.7.3. Problems with Cutoffs and How to Avoid Them

- cutoff introduces a discontinuity in both the potential energy and the force near the cutoff value.
- This creates problems in MD simulations where energy conservation is required. To counteract this discontinuity, one approach is to use a switching function.
- Switching function is a polynomial function of the distance by which the true potential  $v(r)$  by  $v'(r) = v(r)S(r)$

$$v'(r) = v(r) \left[ 1 - 2 \left( \frac{r}{r_c} \right)^2 + \left( \frac{r}{r_c} \right)^4 \right]$$

- Switching function has value of 1 at  $r=0$  and value of 0 at  $r = r_c$ , the cutoff distance
- Switching function prevent simulation from ‘blowing up’

## 6.8. Long-range Forces

- Charge-charge interaction, decays as  $1/r$  is problematic in MD simulations



- A proper treatment of long-range forces can also be important when calculating certain properties, such as dielectric constant.
- One way to tackle inadequate treatment of long-range forces would be to use a much larger simulation cell, which is rather impractical.

### 6.8.1. The Ewald Summation Method

- The Ewald sum was first devised to study the energetics of ionic crystals.
- a particle interacts with all the other particles in the simulation box and with all of their images in an infinite array of periodic cells.
- The charge-charge contribution to the potential energy is due to all pairs of charges in the central simulation box can be written:

$$V = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (6.49)$$

$r_{ij}$  is the minimum distance between charges  $i$  and  $j$

- Ewald sum is the most ‘correct’ way yet devised to accurately include all the effects of long-range forces in computer simulation. Extensively used in simulation involving highly charged systems (ionic melts and in studies of process in and on solid) and systems where electrostatics are important such as lipid bilayer, proteins, and DNA.
- VERY computationally expensive to implement
- electrostatic interactions and simulations using PME is often much more stable, with trajectories much closer to the experimental structures.

$$V = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \left\{ \frac{1}{2} \right.$$

$$V = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \left\{ \begin{aligned} & \sum_{|n|=0}^{\infty} \frac{q_i q_j}{4\pi\epsilon_0} \frac{\text{erfc}(\alpha|r_{ij}+n|)}{|r_{ij}+n|} & (6.55) \\ & + \sum_{k \neq 0} \frac{1}{\pi L^3} \frac{q_i q_j}{4\pi\epsilon_0} \frac{4\pi^2}{k^2} \exp\left(-\frac{k^2}{4\alpha^2}\right) \cos(k \cdot r_{ij}) & (6.57) \\ & - \frac{\alpha}{\sqrt{\pi}} \sum_{k=1}^N \frac{q_k^2}{4\pi\epsilon_0} + \frac{2\pi}{3L^2} \left| \sum_{k=1}^N \frac{q_k}{4\pi\epsilon_0} r_k \right|^2 & (6.58-59) \end{aligned} \right. \quad (6.60)$$