

What MD can do?

MD: interpret experiment, make experimentally testable predictions, and gain a physical understanding of real systems at the atomic level. Sometimes, it shows unexpected results and help to develop a better theory, better understanding, and new experiments.

Properties we can calculate/study:

Pair distribution function in a liquid	Shock waves
Diffusion constant in a liquid	Detonation
Structure factor in a liquid	Radiation damage events
Pressure vs. volume	Grain boundary structure
Thermal expansion	Vapor deposition (*)
Melting point	Fast fracture
Surface melting	Nanoindentation
Free energy differences between phases	Fast plastic flow
Defect dynamics and diffusion in solid (*)	Grain boundary sliding (*)
Surface diffusion (*)	Protein structure (*)
Surface sputtering	Protein function (*)
Surface scattering	Protein dynamics (*)
Thermal conductivity	

(* = may be limited by MD time scale)

Determining Parameters

experimental data

ab initio results

- | | |
|--|--|
| <ul style="list-style-type: none"> • X-ray and neutron scattering crystal structures • vibrational frequencies (IR-Raman) • NMR measurements • crystal lattice constants | <ul style="list-style-type: none"> • Hessian matrix elements → normal modes • forces • energy barriers • electrostatic potential |
|--|--|

Steps in a Typical MD Simulation

1. Prepare molecule
 - Read in pdb and psf file
2. Minimization
 - Reconcile observed structure with force field used (T = 0)
3. Heating
 - Raise temperature of the system
4. Equilibration
 - Ensure system is stable
5. Dynamics
 - Simulate under desired conditions (NVE, NpT, etc)
 - Collect your data
6. Analysis
 - Collect your data
 - Evaluate observables (macroscopic level properties)
 - Or relate to single molecule experiments

Basic idea of MD

- Simulate motions or dynamics of a molecule numerically by computing a trajectory for the system as a function of time: a virtual experiment.
- Basic idea is that we numerically solve Newton Equation

$$F_i = m_i \frac{d^2 x_i}{dt^2}$$

$$x(t) \rightarrow F(x(t)) \rightarrow v(t + \frac{\Delta t}{2}) \rightarrow x(t + \Delta t)$$

Basin MD algorithm

1. Input initial conditions

Potential interaction V as a function of atom positions
Positions r of all atoms in the system
Velocities v of all atoms in the system

2. Compute forces

The force on any atom

$$F_i = -\frac{\partial V}{\partial r_i}$$

is computed by calculating the force between non-bonded atom pairs:

$$F_i = \sum_j F_{ij}$$

plus the forces due to bonded interactions (which may depend on 1, 2, 3, or 4 atoms), plus restraining and/or external forces.

The potential and kinetic energies and the pressure tensor are computed.

3. Update configuration

The movement of the atoms is simulated by numerically solving Newton's equations of motion

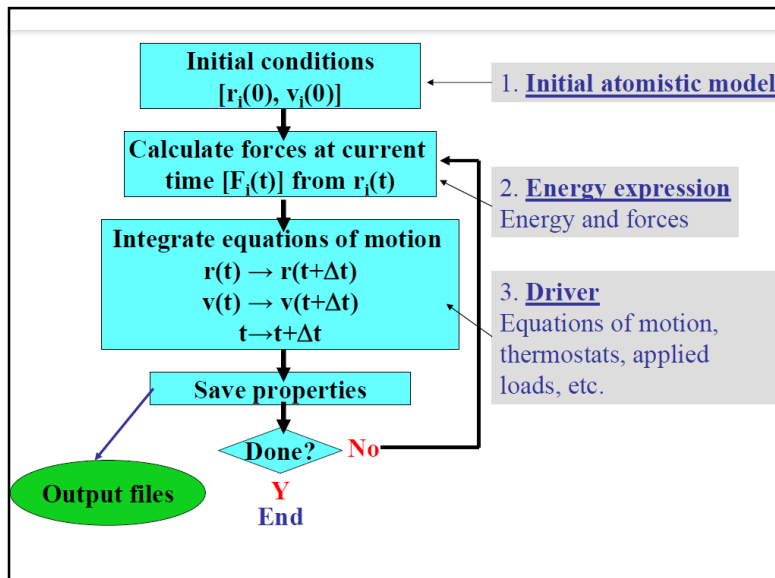
$$\frac{d^2 r_i}{dt^2} = \frac{F_i}{m_i} \quad \text{or} \quad \frac{dr_i}{dt} = v_i; \quad \frac{dv_i}{dt} = \frac{F_i}{m_i}$$

4. if required: **Output step**
write positions, velocities, energies, temperature, pressure, etc.

repeat 2,3,4 for the required number of steps.

Some points of MD

- Potential/force usually have simple function form, although they originate from electronic interaction (electronic wavefunction).
- Born-Oppenheimer approximation.
- Periodic Boundary condition.
- Couple to the heat bath, so that allow the instantaneous T fluctuate around the desired temperature.
- The detailed atomic motions are usually unimportant. What really matters are "the ensemble average" properties. (MD is in fact chaotic with sensitive dependence on initial conditions)



Integrating the equations of motion

$$\left. \begin{aligned} \dot{\vec{r}}_i &= \frac{\vec{p}_i}{m_i} \approx \frac{\vec{r}_i(t + \Delta t) - \vec{r}_i(t)}{\Delta t} & \vec{r}_i(t + \Delta t) &= \vec{r}_i(t) + \frac{\vec{p}_i}{m_i} \Delta t \\ \dot{\vec{p}}_i &= \vec{F}_i \approx \frac{\vec{p}_i(t + \Delta t) - \vec{p}_i(t)}{\Delta t} & \vec{p}_i(t + \Delta t) &= \vec{p}_i(t) + \vec{F}_i \Delta t \end{aligned} \right\} \text{Euler method}$$

Verlet algorithm: Taylor expansion of positions with time

$$r_i(t + \Delta t) = r_i(t) + \dot{r}_i(t)\Delta t + \frac{1}{2}\ddot{r}_i(t)\Delta t^2 + \frac{1}{6}\dddot{r}_i(t)\Delta t^3 + O(\Delta t^4)$$

$$r_i(t - \Delta t) = r_i(t) - \dot{r}_i(t)\Delta t + \frac{1}{2}\ddot{r}_i(t)\Delta t^2 - \frac{1}{6}\dddot{r}_i(t)\Delta t^3 + O(\Delta t^4)$$

Sum two equations:

$$r_i(t + \Delta t) = 2r_i(t) - r_i(t - \Delta t) + \ddot{r}_i(t)\Delta t^2 + O(\Delta t^4)$$

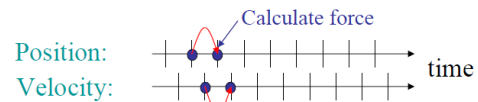
Verlet algorithm: various equivalent formulations

Velocity Verlet:

$$\begin{cases} p_i\left(t + \frac{1}{2}\Delta t\right) = p_i(t) + \frac{1}{2}F_i(t)\Delta t & \text{Advance } p \text{ half step} \\ r_i(t + \Delta t) = r_i(t) + \frac{p_i\left(t + \frac{1}{2}\Delta t\right)}{m_i}\Delta t & \text{Advance } r \text{ full step (with } p \text{ half step ahead)} \\ p_i(t + \Delta t) = p_i\left(t + \frac{1}{2}\Delta t\right) + \frac{1}{2}F_i(t + \Delta t)\Delta t & \text{Advance } p \text{ half step (with } F \text{ at } t + \Delta t) \end{cases}$$

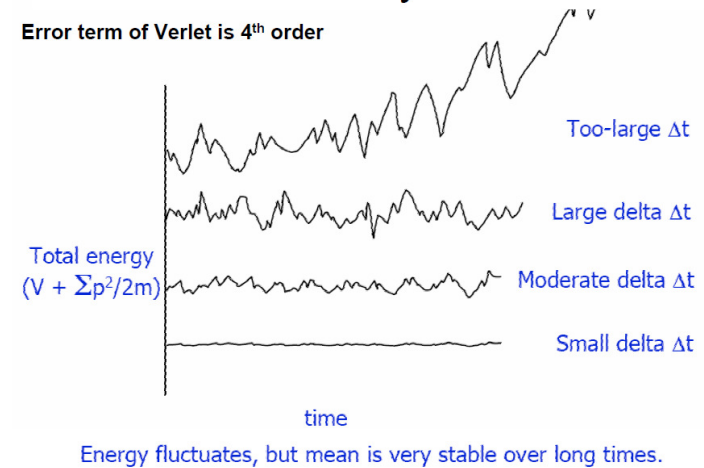
Leap-Frog Verlet:

$$\begin{cases} r_i\left(t + \frac{1}{2}\Delta t\right) = r_i\left(t - \frac{1}{2}\Delta t\right) + v_i(t)\Delta t \\ v_i(t + \Delta t) = v_i(t) + \frac{1}{m_i}F_i\left(t + \frac{1}{2}\Delta t\right) \end{cases}$$



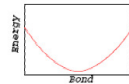
Accuracy Vs. Δt

Error term of Verlet is 4th order



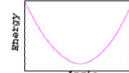
Commonly used potentials:

Bonds



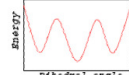
$$v_{bond} = \frac{1}{2}k_b(d_i - d_i^0)^2$$

Angles



$$v_{angle} = \frac{1}{2}k_\theta(\theta_i - \theta_i^0)^2$$

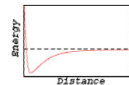
Dihedrals



$$v_{dihedral} = k_\phi[1 + \cos(3\phi_i)] + k'_\phi[1 - \cos(\phi_i - \pi)]$$

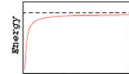
From quantum chemistry, thermodynamics

Van der Waals



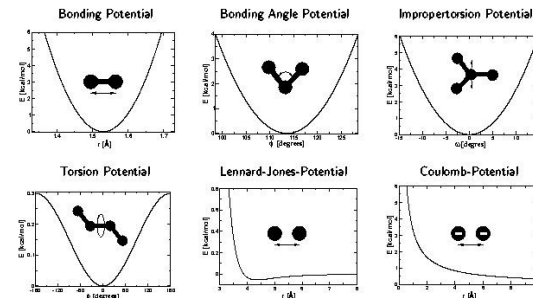
$$v_{vdW} = \epsilon_{ij}^{min} \left[\left(\frac{d_{ij}^{min}}{d_{ij}} \right)^{12} - 2 \left(\frac{d_{ij}^{min}}{d_{ij}} \right)^6 \right]$$

Electrostatics



$$v_{Coulomb} = \frac{q_i q_j}{\epsilon d_{ij}}$$

Molecular Mechanics Force Field



CHARMM Energy Function:

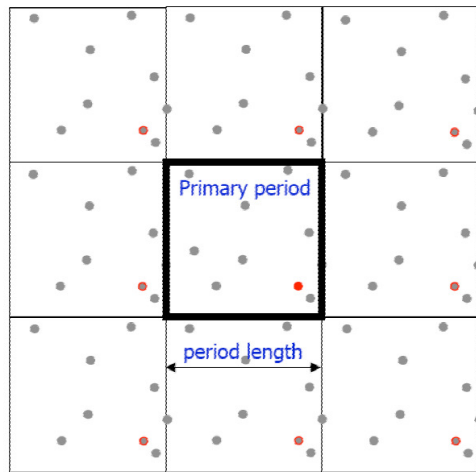
$$V(R) = E_{bonded} + E_{non-bonded}$$

$$E_{bonded} = \sum_{bonds} k_b(b - b_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} k_\phi(1 + \cos[n\phi - \delta]) + \sum_{impropers} k_\omega(\omega - \omega_0)^2$$

$$E_{non-bonded} = \sum_{i,j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{i,j} \frac{1}{4\pi\epsilon} \frac{q_i q_j}{r_{ij}}$$

Periodic boundary conditions

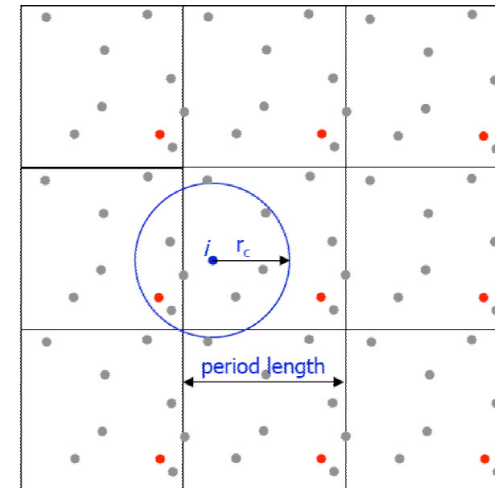
- To minimize edge effects in a finite system, apply the periodic boundary condition;
- "Periodic Boundary Conditions" means that the box is surrounded by copies of itself.



Minimum-image condition

If $r_c < \text{period length}/2$, then only one image of each neighbor will be within cutoff distance of atom i .

This simplifies construction of the neighbor lists, and reduces artificial behavior.



Temperature and thermostat

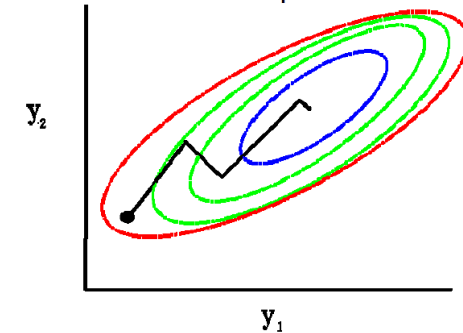
$$T = \sum \frac{1}{2} m (\vec{v}_i - \vec{v}_c)^2 / (\frac{3}{2} N k_B)$$

- Thermostat: the system may be coupled to a heat bath to ensure that the average system temperature is maintained close to the requested temperature
- Berendsen thermostat: rescale all velocity by a factor of λ to control T .

$$\lambda = [1 + \frac{\Delta t}{\tau} (\frac{T_{desire}}{T} - 1)]^{1/2}$$

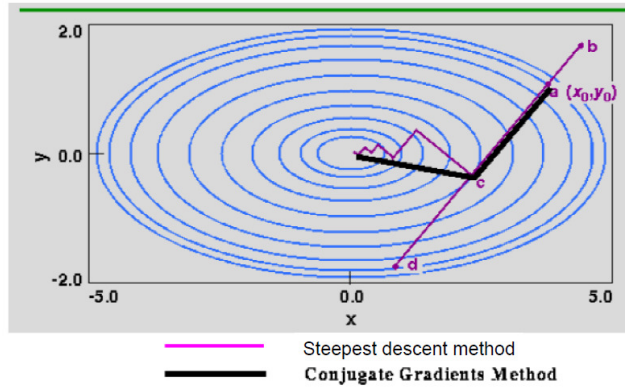
Energy minimize: steepest descent

- steepest descents: we move the molecule along the force vector reducing the energy as much as possible in each step

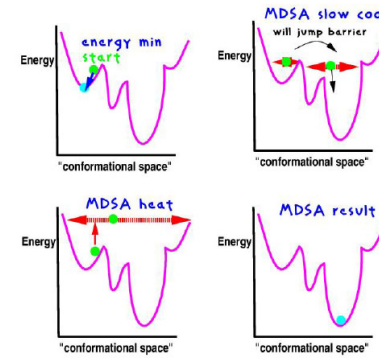


Energy minimize: Conjugate gradient

- Conjugate gradient is efficient compared to steepest descents;



Energy minimize: Simulated Annealing



We have a new approach: basin-filling + simulated annealing to search for the global minimization!

Global minimization: Monte Carlo Simulated Annealing, Molecular Dynamics Simulated Annealing and Genetic Algorithms are the most popular approaches to provide the partial solution.

Pair distribution function

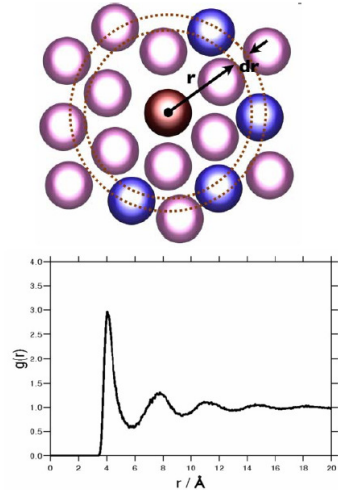
$$g(\vec{r}) = \frac{V}{N^2} \left\langle \sum_i \sum_{j \neq i} \delta(\vec{r} - \vec{r}_{ij}) \right\rangle$$

$$g(r) = \int g(\vec{r}) \sin \theta d\theta d\phi$$

Structure factor:

$$S(\vec{q}) = \int g(\vec{r}) e^{i\vec{q} \cdot \vec{r}} d\vec{r}$$

X-ray scattering intensity: $I(\vec{q}) = |S(\vec{q})|^2$



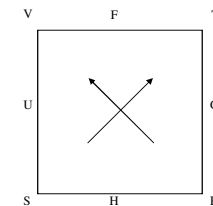
Some common ensembles used in molecular simulations:

Microcanonical ensemble (NVE) – maximum entropy

Canonical ensemble (NVT) – minimum Helmholtz free energy

Isothermal-isobaric ensemble (NPT) – minimum Gibbs free energy

Grand canonical ensemble (μVT) – maximum PV



Microcanonical (NVE)	Canonical (NVT)	Isobaric/isothermal (NPT)
Probability distributions		
$P(\{r_i\}, \{p_i\}) = \frac{1}{\Omega(E, V, N)}$	$P(\{r_i\}, \{p_i\}) = \frac{e^{-\frac{H(\{r_i\}, \{p_i\})}{kT}}}{Z(T, V, N)}$	$P(\{r_i\}, \{p_i\}, V) = \frac{e^{-\frac{1}{kT}[H(\{r_i\}, \{p_i\}) - PV]}}{Z_P(T, P, N)}$
$\Omega(E, V, N) = \sum_{micro} \delta(E - H(\{r_i\}, \{p_i\}))$	$Z(T, V, N) = \sum_{micro} e^{-\frac{E}{kT}}$	$Z_P(T, P, N) = \sum_V \sum_{micro} e^{-\frac{E - PV}{kT}}$
Free energies		
$S = k \log \Omega(E, V, N)$	$F(T, V, N) = -kT \log Z$	$G(T, P, N) = -kT \log Z_P$

Some simple thermodynamic quantities:

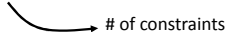
internal energy: $U = \langle E \rangle = \frac{1}{M} \sum_{i=1}^M E_i$

heat capacity: $c_v = \left(\frac{\partial U}{\partial T} \right)_v$

$$c_v = \left(\frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2} \right) = \left(\frac{\langle (E - \langle E \rangle)^2 \rangle}{k_B T^2} \right)$$

temperature:

$$KE = \sum_{i=1}^N \frac{|\mathbf{p}_i|^2}{2m_i} = \frac{k_B T}{2} (3N - N_c)$$

 # of constraints