# MATE 664 Lecture 22

## Multiscale Simulation for Kinetics of Materials (III): Molecular Dynamics (MD)

Dr. Tian Tian

2026-03-25

### Learning outcomes

After this lecture, you will be able to:

- **Recall** the basic idea of molecular dynamics as time integration of atomic motion
- **Identify** the main components of an MD simulation, including forces, integrators, and ensembles
- **Describe** how thermostats and timesteps affect MD simulations
- **Interpret** the role of interatomic potentials in molecular dynamics

### Announcement: class presentation

- **Date**: April 8 14:00 - 17:30
- **Location**: HC 2-14 (same classroom)
- **Format**: slide presentation or your choice of media. Pre-recorded video allowed if unable to attend
- **Length**: 10 min presentation + 2 min questions

### Recap: Monte Carlo (MC) and Kinetic Monte Carlo (KMC) methods

- (Metropolis) MC: sampling equilibrium configuration
- KMC: sampling configuration change over (pseudo) time scale
- Major assumption: average / integral of states can be performed by random sampling
- Probability ratio (NVT)

$$p_i \propto e^{-E_i/k_B T} \tag{1}$$

## Recap: potential energy comparison in MC

The central parameter in MC / KMC simulations is the estimation of probability, or equivalent the potential energy $E_i$ (or $U_i$).

Remaining questions:

- where do we get these energy values?
- if we do need to calculate force $-\frac{\partial E}{\partial x}$, what can we get?
- can we get continuous motion instead of discrete states "hopping"?

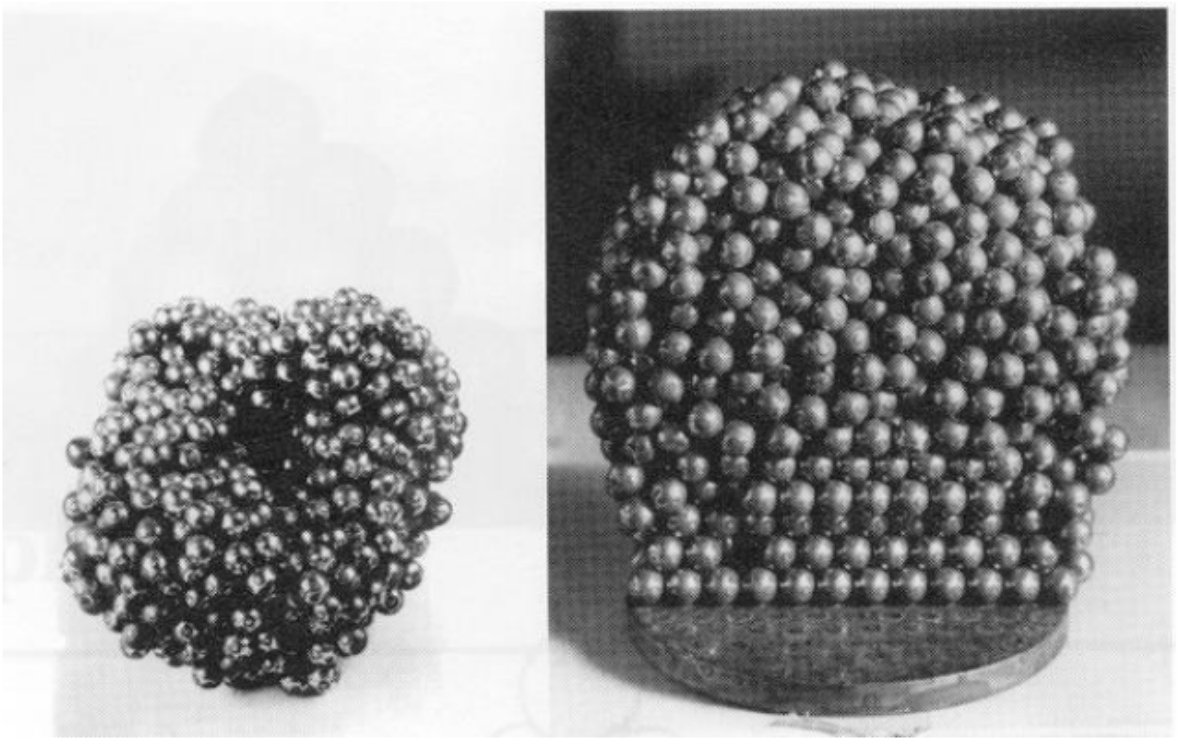
## Historical review: atomic hypothesis

Excerpt from Richard Feynman's physics lecture at Caltech

> I believe it is the atomic hypothesis (or the atomic fact, or whatever you want to call it) that all things are made out of atoms – little particles that move around, are in perpetual motion, attract each other when they are some distance apart, but repel being squeezed into one another. In that one sentence you'll see there's an enormous amount of information about the world if just a little imagination and thinking is applied.

## Representing large system with small domain

Pioneered by Bernal (UCL) in the 1950s, use ball bearing to study the liquid-solid phase transformation.



## Hard-sphere models for materials simulations

Simulating defects with ball bearings!

https://www.youtube.com/watch?v=O3RsDIWB7s0

## Another way down the road: molecular dynamics (MD)

Instead of studying the systems in configuration space ($\{\mathbf{r}_i, \cdot\}$), MD studies materials in a phase space ($(\{\mathbf{r}_i, \mathbf{p}_i), \cdot\}$), more naturally representing the Newton's equation

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m_i} \tag{2}$$

$$\frac{d\mathbf{p}_i}{dt} = -\nabla_{r_i} U(\mathbf{r}) = \mathbf{F}_i \tag{3}$$

Time dependency now appears!

## MD = energy + dynamics

### Energy

- potential energy surface (PES)
- bonding physics
- force evaluation
- material specificity

### Dynamics

- Newton's equations
- time integrator
- thermostat / barostat
- trajectory sampling

- no MD without both parts
- wrong energy or wrong dynamics both lead to wrong physics (butterfly effect)

## Dissecting the MD method

- **Deterministic** method: state of the system at any future time can be predicted from its current state (with caution)
- MD cycle for one step:
  - Force acting on each atom is assumed to be constant during the time interval
  - Forces on the atoms are computed and combined with the current positions and velocities to generate new positions and velocities a short time ahead
- MD simulations provide a molecular level picture of structure and dynamics → structure-property relationships (**SPR**)

## Timescale in MD

### Physical world

- Bond vibrations - 1 fs
- Collective vibrations - 1 ps
- Conformational transitions - ps or longer
- Enzyme catalysis - microsecond/millisecond
- Ligand Binding - micro/millisecond
- Protein Folding - millisecond/second

**MD Simulation**

- Integration time step - 1 femtosecond
- Set by fastest varying force
- Accessible timescale about **10 nanoseconds**

## Topic 1: the dynamics

We wanted to talk about the "control" part of the MD

- how do we turn forces into trajectories?
- how do we know the simulation system represents the set ensemble?
- how to get quantities from MD simulation?

## A bit of statistical mechanics: averaging quantity in MD

In MD simulation a macroscopic quantity is averaged over time

$$\langle A \rangle_{\text{time}} = \frac{1}{\tau} \int_0^\tau A(t)\, dt \tag{4}$$

The ergodic assumption gives that for an infinitely long enough simulation, every configuration in the phase space should eventually be visited. (Like MC, but we evolve continuously in time)

$$\langle A \rangle_{\text{ens}} = \int d\mathbf{r}^N d\mathbf{p}^N\, A(\mathbf{r}^N, \mathbf{p}^N)\, \rho(\mathbf{r}^N, \mathbf{p}^N) \tag{5}$$

**Time average = Ensemble average**

## What quantities are we interested in MD?

What quantities are we interested in MD?

Thermodynamic

- temperature $T$
- pressure $P$
- energy $E$

Structural

- radial distribution function $g(r)$
- coordination number $N_c$
- defect structure / order parameter $c_i(\mathbf{r})$

**Kinetic**

- **diffusivity** $D$
- viscosity $\eta$
- relaxation time $\tau$
- conductivity $\mu$

## Ensembles: how to describe your system?

In a MD setup, we usually want to fix a few macroscopic quantities while let others to give our statistical average

Relate to thermodynamic relations

$$dU = T\,dS - P\,dV + \mu\,dN$$

- common MD ensembles:
  - NVE (micro-canonical)
  - NVT (canonical)
  - NPT (isothermal- isobaric)

## Comparison between ensembles

NVE

- isolated system
- fixed $N, V, E$
- Probability $\rho_{NVE} \propto \delta(H - E)$

NVT

- heat bath
- fixed $N, V, T$
- $\rho_{NVT} \propto e^{-H/k_B T}$

NPT

- heat + pressure bath
- fixed $N, P, T$
- $\rho_{NPT} \propto e^{-(H+PV)/k_B T}$

## MD: how to advance in time?

Just solve the coupled dynamic equations

$$\dot{\mathbf{r}}_i = \mathbf{v}_i \tag{6}$$
$$m_i \dot{\mathbf{v}}_i = \mathbf{F}_i \tag{7}$$

At each time step:

1. compute forces
2. update positions and velocities
3. apply ensemble control if needed
4. collect observables

## The Verlet algorithm

We can of course use $t \to t + \Delta t$ (the forward Euler) method, but maybe not so stable.

Verlet algorithm: use Taylor expansion

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \mathbf{a}_i(t)\Delta t^2 + O(\Delta t^4) \tag{8}$$

Requires 2 time steps to advance

$$\mathbf{v}_i(t) = \frac{\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t)}{2\Delta t} \tag{9}$$

Can now go back in time!

## Velocity Verlet: most common practical form

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\Delta t + \frac{1}{2}\mathbf{a}_i(t)\Delta t^2 \tag{10}$$

$$\mathbf{v}_i\left(t+\frac{\Delta t}{2}\right) = \mathbf{v}_i(t) + \frac{1}{2}\mathbf{a}_i(t)\Delta t \tag{11}$$

$$\mathbf{a}_i(t+\Delta t) = \frac{\mathbf{F}_i(t+\Delta t)}{m_i} \tag{12}$$

$$\mathbf{v}_i(t+\Delta t) = \mathbf{v}_i\left(t+\frac{\Delta t}{2}\right) + \frac{1}{2}\mathbf{a}_i(t+\Delta t)\Delta t \tag{13}$$

## MD: choice timestep

$\Delta t$ must resolve the fastest motion

**too large**:

- unstable integration
- broken energy conservation

**too small**:

- too expensive

## MD: typical values for timesteps

- metal MD: $\sim 1$ fs
- molecules with constrained bonds: $\sim 2$ fs
- coarse motion / rigid modes: up to $\sim 10$ fs
- *ab initio* MD: often smaller than 1 fs

## MD: how to maintain ensemble?

- **NVE**: integrate Newton's law only
- **NVT**: add thermostat
- **NPT**: thermostat + barostat

Ideas for thermo- and barostats:

- thermostat controls kinetic energy
- barostat allows box size / shape to fluctuate

## Comparison between thermostats

Berendsen

- fast equilibration
- weakly rescales velocities
- does not generate exact canonical ensemble

Nosé-Hoover

- extended dynamical system
- samples canonical ensemble better
- may oscillate if poorly tuned

Langevin

- friction + random force
- good temperature control
- can distort dynamics if damping is too strong

## Topic 2: the energy

- How do you get the energy and forces?
- Which energy form is good?
- Can we get energy from data?

## Potential energy in MD (1)

### Interatomic Potential

- Ab initio, electronic structure methods
  - Hartree-Fock, Density Functional Theory
  - Quantum Monte Carlo
  - Very Accurate but computationally intensive (applicable to small systems – 100s-1000s atoms)
- Interatomic potentials:

$$V(\{r_i\}) = V_{\text{cov/met}}(\{r_i\}) + V_{electr}(\{r_i\}) + V_{vdW}(\{r_i\})$$

Parameterized to describe specific materials

## Potential energy in MD (2)

### Energy and Force

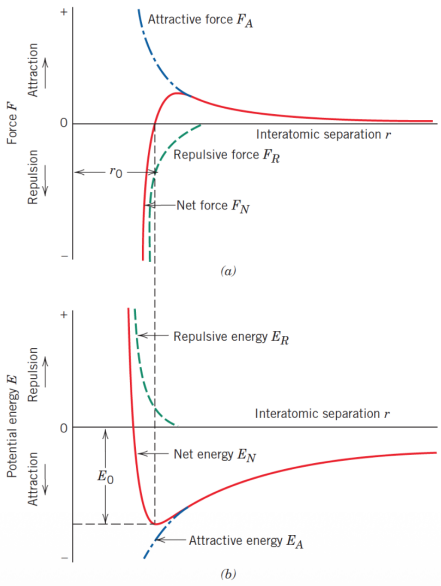
$F_N=F_A+F_R$
$r_0$: Equilibrium distance
$F_N=0$: state of equilibrium
$F_N>0$: attractive
$F_N<0$: repulsive

It is more convenient to use potential energy

$$U = \int F\,dr$$

$$U_N = \int_r^\infty F_N\,dr = \int_r^\infty F_A\,dr + \int_r^\infty F_R\,dr = U_A + U_R$$

$r_0$→ lowest energy
$E_0$ or $\varepsilon$→ bond energy; it is the energy that requires separating two atoms to an infinite distance.



## Potential energy in MD (3)

### EAM

Embedded atom model (EAM)

Mike Baskes and collaborators 1980's + other groups

$$V = \sum_{i<j} \phi(r_{ij}) + \sum_{i} F(\rho_i) \qquad \rho_i = \sum_{j \neq i} f(r_{ij})$$

Two-body potential

Embedding energy: Energy to embed atom *i* in the electronic density $\rho_i$

Local electronic density around atoms i contributed from nearby atoms

**Accurate description of:**

- Environment dependence of bonding in metals (vacancy and surfaces)
- Elastic constants and plastic deformation
- Phase transformations (melting & solid-solid)
- Alloys
- Fit parameters based on experimental data
  - Lattice parameter, elastic constants, vacancy & surface energy
- Use of *ab initio* data enables the incorporation of more information

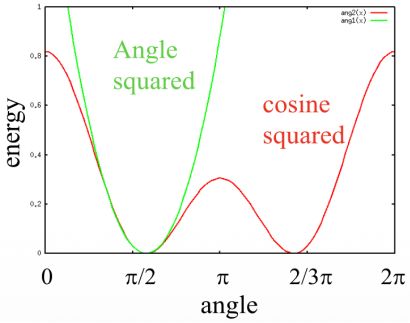
## Potential energy in MD (4)

### Covalent Interaction

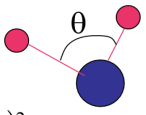
Interaction between atoms separated by chemical bonds

Bond stretch:

$$\phi_{bond}(r) = \begin{cases} \frac{1}{2} k_{ij} \left(r - r_{ij}^0\right)^2 \\ Morse, L-J \end{cases}$$




Angle bending:




$$\phi_{angle}(\theta) = \begin{cases} \frac{1}{2} k_{ijk} \left(\theta - \theta_{ijk}^0\right)^2 \\ \frac{1}{2} \frac{k_{ijk}}{\sin^2 \theta_{ijk}^0} \left[\cos\theta - \cos\theta_{ijk}^0\right]^2 \end{cases}$$

## More about the atomic hypothesis

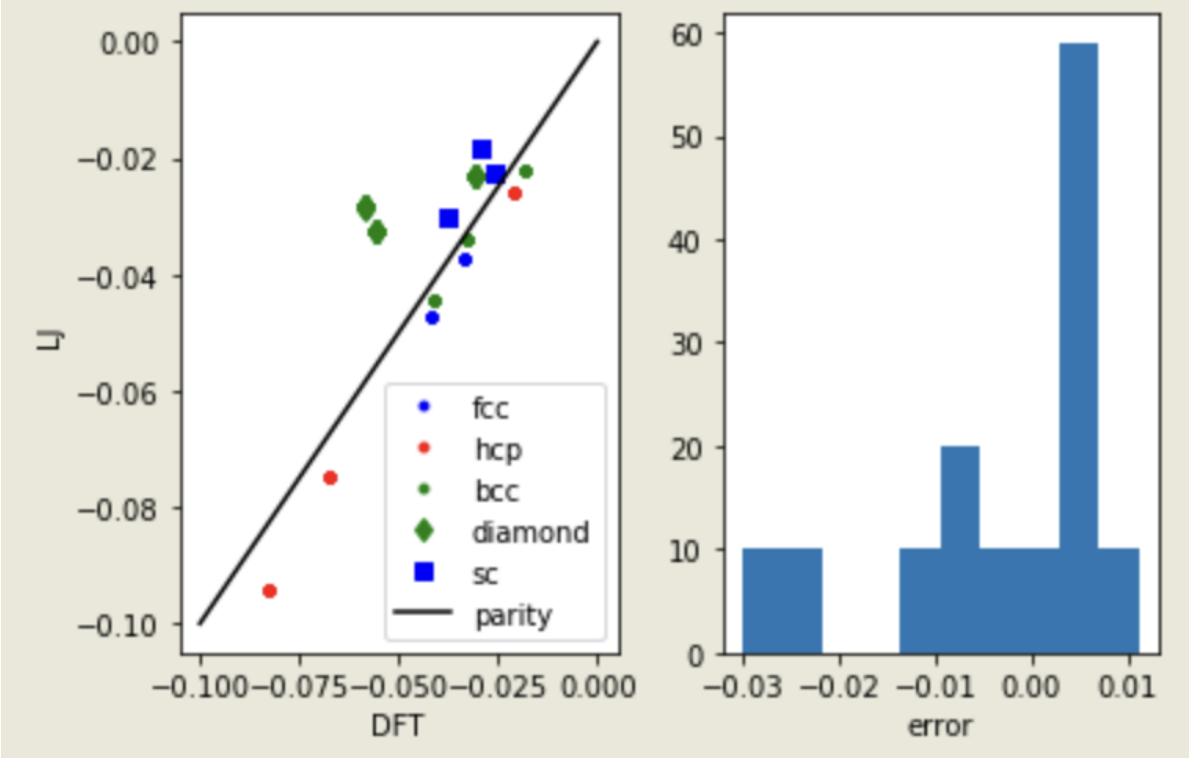
Still from Feynman:

> If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed onto the next generations of creatures, what sentence would contain the most information in the fewest words? I believe it is the atomic hypothesis that all things are made of atoms – little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon squeezed into one another. In that sentence, you will see, there is enormous amount of information about the world, if just a little imagination and thinking are applied

## Example of data-driven potential

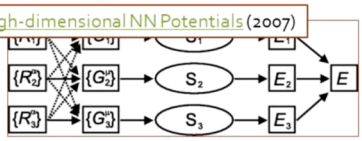
Can we use data to train potential? Examples from John Kitchin (CMU)

- Atomic Simulation Environment (ASE) as a easy python wrapper over the atomistic calculator
- Bulk argon energies calculated in different lattice type, volume etc.
- Fit the polynomial LJ potential using the DFT dataset!

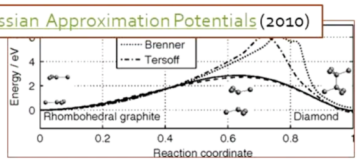



## Interatomic potentials: from polynomials to machine learning

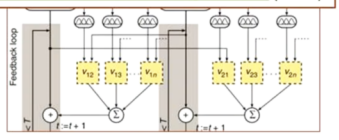
High-dimensional NN Potentials (2007)



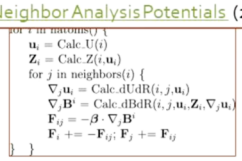

Gaussian Approximation Potentials (2010)




Deep Tensor Neural Networks / MPNNs (2016)




Spectral Neighbor Analysis Potentials (2014)




Moment Tensor Potentials (2015)


MOMENT TENSOR POTENTIALS: A CLASS OF SYSTEMATICALLY IMPROVABLE INTERATOMIC POTENTIALS*

ALEXANDER V. SHAPEEV†

And many other works:
ANI-1 (2017), AGNI (2015), MAISE (2009),
Sparse Interatomic Potentials (2014), MBTR (2017),
Genetic Programming (2020),

## Summary

- Molecular dynamics evolves atomic positions and momenta using Newton's equations
- MD simulations require interatomic forces, numerical integration, and ensemble control
- The choice of potential strongly affects what material behavior an MD model can represent