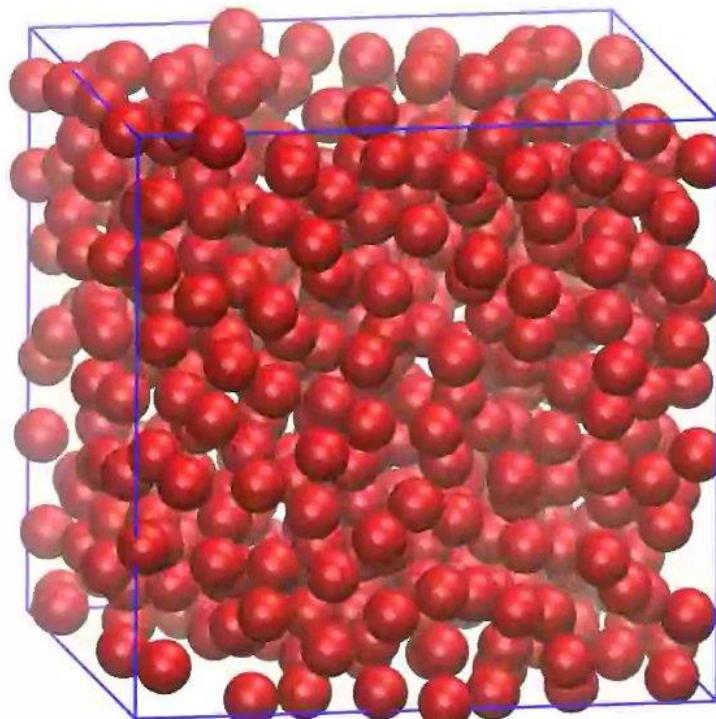Physics 5403: Computational Physics

Chapter 6: Molecular Dynamics

6.0 Overview

What is molecular dynamics (MD)?

Numerical method for studying many-particle systems such as molecules, clusters, and even macroscopic systems such as gases, liquids and solids

Used extensively in materials science, chemical physics, and biophysics/biochemistry
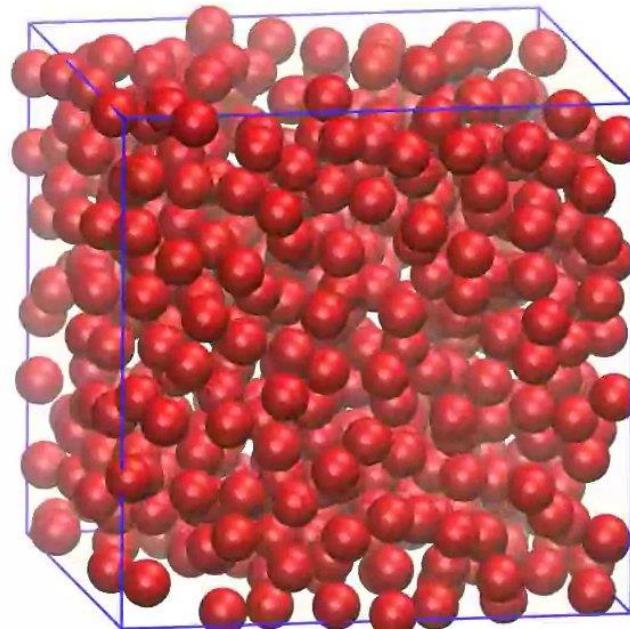
First reported MD simulation:

Alder + Wainwright (1957): Phase diagram of a hard-sphere gas

Basic idea of molecular dynamics:

Solution of Newton's equations of motion for the individual particles (atoms, ions, …)

Example:

Molecular dynamics simulation of liquid argon

Advantages:

- gives (in principle) **complete knowledge** of system; if all trajectories are known, **everything** can be calculated

- easily accommodates **non-equilibrium states** and other complex situations beyond thermal equilibrium (by preparing appropriate initial conditions)

Disadvantages:

- complete knowledge of all trajectories is **often much more information** than needed (e.g., equilibrium state of a fluid is characterized by just two variables, p and T)

- Is this approach **efficient**?

Classical MD vs. ab-initio MD

Questions:

(i)   Is the classical description of the particles in terms of Newtonian mechanics justified?

(ii) What are the forces between the particles; how can we determine them?

Classical vs. quantum description:

- compare inter-particle distance to de-Broglie wavelength

$$\lambda = h/p \qquad E_{kin} = \frac{p^2}{2m} \approx \frac{3}{2} k_B T$$

$$\lambda = \frac{h}{\sqrt{3m\, k_B T}}$$

argon $\quad m = 40\, m_p \quad, \quad T = 300\, K$

$$\lambda \approx 0.1\, \text{Å} \quad \ll \quad 4\, \text{Å} \quad (\text{typical distances}).$$

→ Motion of atoms and ions is **classical**

Interaction potentials and forces:

- interaction between atoms and molecules results from electronic structure: not a classical problem, **requires quantum** physics

- two different ways to proceed, leading to two different classes of molecular dynamics simulations, classical MD and ab-initio MD

Classical molecular dynamics

- Interactions are approximated by classical **model potentials** constructed by comparison with experiment (empirical potentials)

- Leads to simulation of purely classical many-particle problem

- Works well for simple particles (such as noble gases) that interact via isotropic pair potentials

- Poor for covalent atoms (directional bonding) and metals (electrons form Fermi gas)

- Simulations **fast**, permit **large particle numbers**

Ab-initio molecular dynamics

- Performs a full quantum calculation of the electronic structure at every time step (for every configuration of the atomic nuclei),

  ab-initio = from first principles

- Forces are found the dependence of the energy on the particle positions

- Much higher accuracy than classical MD, but much higher numerical effort (restricts number of particles and simulation time)

In this course:

- Classical molecular dynamics only

Resources:

- MD Primer by Furio Ercolessi

    www.fisica.uniud.it/~ercolessi

- Molecular dynamics codes:

    LAMMPS,

    GROMACS

    + several others

6.1 Basic Machinery

N particles at positions $\vec{r}_i$ (i=1...N), velocities $\dot{\vec{r}}_i = \frac{\vec{p}_i}{m}$

Kinetic energy:
$$E_{kin} = \sum_{i=1}^{N} \frac{m_i}{2} \dot{\vec{r}}_i^2 = \sum_{i=1}^{N} \frac{\vec{p}_i^2}{2 m_i}$$

Potential energy:
$$E_{pot} = V_{int}(\vec{r}_1, \ldots, \vec{r}_N) + \sum_{i=1}^{N} U_{ext}^{(i)}(\vec{r}_i)$$

$\underbrace{\phantom{V_{int}(\vec{r}_1, \ldots, \vec{r}_N)}}_{\text{interaction between particles}}$ $\underbrace{\phantom{\sum_{i=1}^{N} U_{ext}^{(i)}}}_{\text{external potential}}$

Forces:
$$\vec{F}_i = -\frac{\partial}{\partial \vec{r}_i} V_{int}(\vec{r}_1, \ldots \vec{r}_N) - \frac{\partial}{\partial \vec{r}_i} U_{ext}^{(i)}(\vec{r}_i)$$

Total energy: $E = E_{kin} + E_{pot}$

The interaction potential $V_{int}(\vec{r}_1, \dots \vec{r}_N)$

Simplest choice: $V_{int}(\vec{r}_1, \dots, \vec{r}_N) = \dfrac{1}{2} \displaystyle\sum_{i \neq j = 1}^{N} V(\vec{r}_i - \vec{r}_j)$

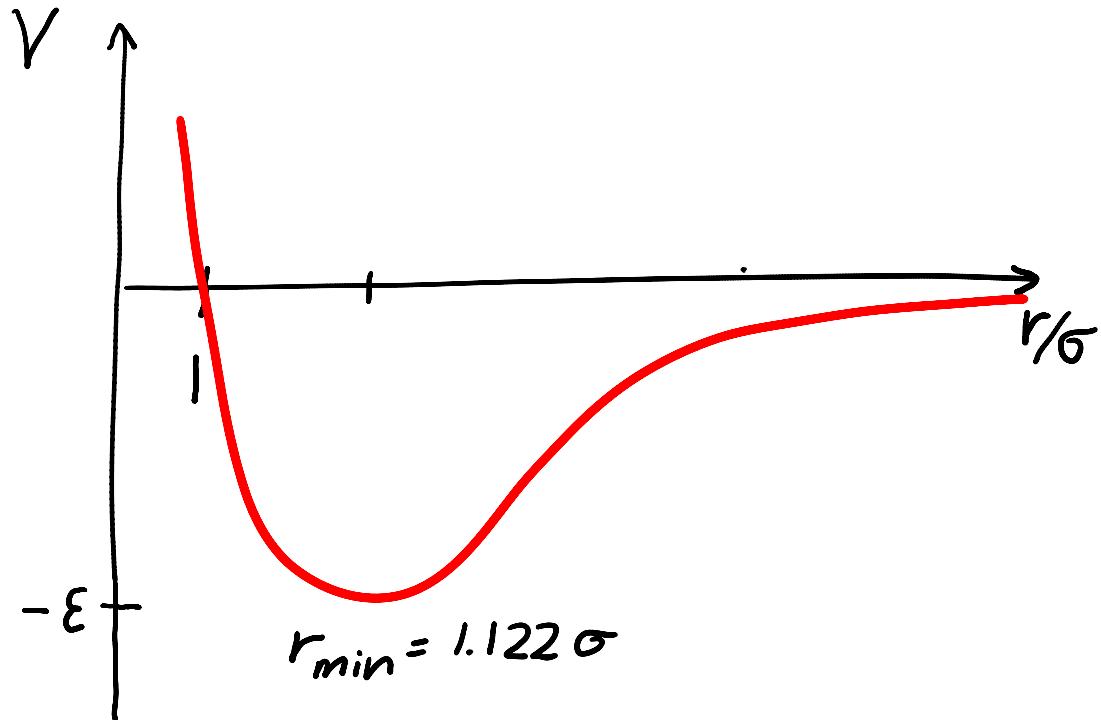Interaction is sum over distance dependent pair potential

Implies:
- No internal degrees of freedom (spherical particles) H_2O
- No directional bonding (like in covalent materials, e.g. semiconductors, carbon)

→ works well for closed-shell atoms like the noble gases

Popular model pair potential: **Lennard-Jones Potential**

$$V(\vec{r_i} - \vec{r_j}) = V(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right]$$

$\frac{1}{r^6}$ term:
Van-der-Waals attraction
between neutral atoms
(induced dipole-dipole)

$\frac{1}{r^{12}}$ term:
Short-range repulsion due to
overlap of the electron clouds;
form purely phenomenological

V

r/σ

$-\varepsilon$

$r_{min} = 1.122\,\sigma$

Truncation:

Lennard-Jones potential goes out to r→∞
One has to calculate a large number of small contributions

Sometimes V(r) will be truncated at R_C V(r)≡0 for r>R_C

To avoid potential jump at R_C: shift

$$V_T(r) = \begin{cases} V(r) - V(R_c) & r \le R_c \\ 0 & r > R_c \end{cases}$$

Common truncation radii for the Lennard-Jones potential are
2.5σ or 3.2σ

6.2 Time integration –Verlet algorithm

Equations of motion are 2nd order ODE for positions $\vec{r}_i$ (i=1,...,N) → in principle one could use any of the integration algorithms discussed in chapter 3 to integrate the equations of motion

However:

Integration can be simplified by making use of the special structure of the equation of motion: forces depend only on $\vec{r}$, not $\dot{\vec{r}}$

$$m_i \frac{d^2}{dt^2} \vec{r}_i = \vec{F}_i(\vec{r}_1,..,\vec{r}_N)$$

Verlet algorithm

$$t_n = n\tau$$

$$\frac{d^2}{dt^2}\vec{r_i} = \vec{G_i}(\vec{r_1},\dots,\vec{r_N}) \qquad\qquad \vec{G_i} = \frac{\vec{F_i}}{m_i}$$

$$\frac{d^2}{dt^2}\vec{r_i}(t_n) = \frac{1}{\tau^2}\left(\vec{r_i}(t_{n+1}) - 2\vec{r_i}(t_n) + \vec{r_i}(t_{n-1})\right) + O(\tau^2)$$

$$\vec{r_i}(t_{n+1}) = 2\vec{r_i}(t_n) - \vec{r_i}(t_{n-1}) + \tau^2\vec{G_i}\left(\{\vec{r}(t_n)\}\right) + O(\tau^4)$$

Verlet algorithm

$$\vec{r}_i(t_{n+1}) = 2\vec{r}_i(t_n) - \vec{r}_i(t_{n-1}) + \tau^2 \vec{G}_i(\{\vec{r}(t_n)\}) + O(\tau^4)$$

- Position error is 4th order in τ (almost like Runge-Kutta, ch. 3)
- Math simpler than two Runge-Kutta algorithms required for a 2nd order ODE

Note: velocities do not show up! If velocities are desired:

$$\dot{\vec{r}}_i(t_n) = \frac{\vec{r}_i(t_{n+1}) - \vec{r}_i(t_{n-1})}{2\tau} + O(\tau^2)$$

Advantages:

- Very simple
- High accuracy for the positions
- If velocities are not needed, their calculation can be skipped

Disadvantages:

- Accuracy of velocities is only $O(\tau^2)$
- Method is not self-starting, $\vec{r}_i(t_n)$ and $\vec{r}_i(t_{n-1})$ are necessary
  Usually $\vec{r}_i(t_0)$ and $\dot{\vec{r}}_i(t_0)$ are given $\rightarrow$ construct

$$\vec{r}_i(t_1) = \vec{r}_i(t_0) + \tau\, \dot{\vec{r}}_i(t_0) + \frac{\tau^2}{2}\, \vec{G}_i\left(\{\vec{r}_i(t_0)\}\right)$$

There are modifications of the Verlet algorithm that treat $\vec{r}$ and $\dot{\vec{r}}$ on equal footing.

6.3 Geometry and boundary conditions

Must distinguish simulation of finite objects like molecules and clusters from simulation of macroscopic systems
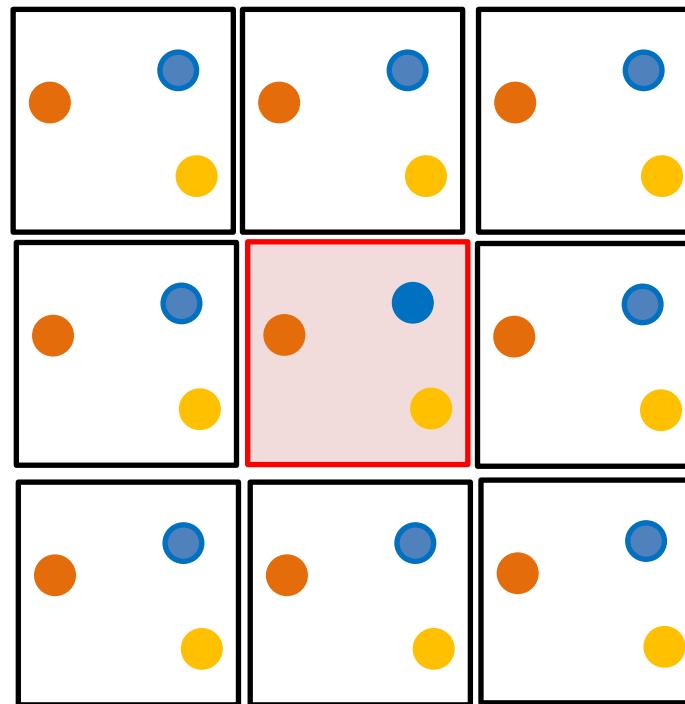
- Finite systems: use open boundary conditions, i.e. no boundaries at all, just N particles in space

- Macroscopic systems: real macroscopic systems have a much larger number of particles ($\sim 10^{23}$) than can be handled in a simulation

$\rightarrow$ simulating a large cluster with open boundary conditions will greatly overestimate surface effects
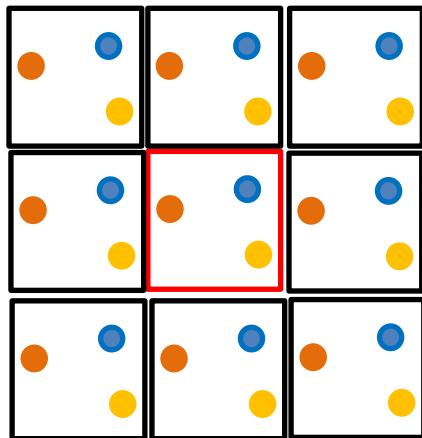
Solution: periodic boundary conditions

Periodic boundary conditions

Consider box of size L, repeat box infinitely many times in all directions

Basic box

copies

Periodic boundary conditions

Consider box of size L, repeat box infinitely many times in all directions

Each particle interacts (in principle) with all particles in all boxes → problems for long-range interactions (infinite resummation necessary)
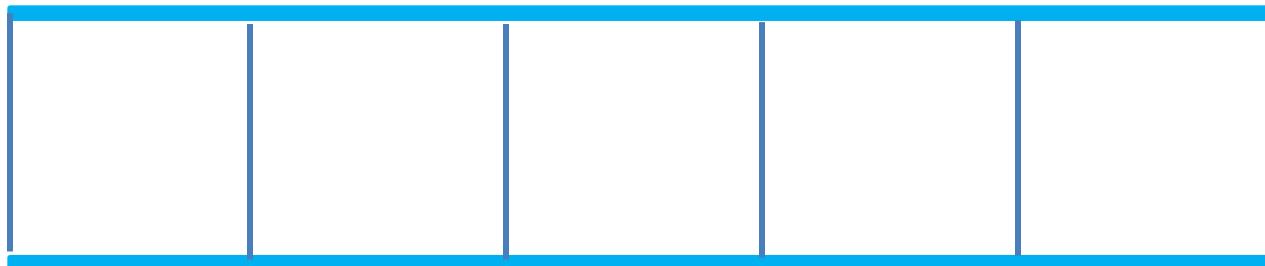
short-range interactions: minimum image convention: consider box with size $L > 2R_C$, at most the closest of all images of a particle j can interact with a given particle i → great simplification: pick the closest image and use this to calculate $V(r_{ij})$

Systems with surfaces
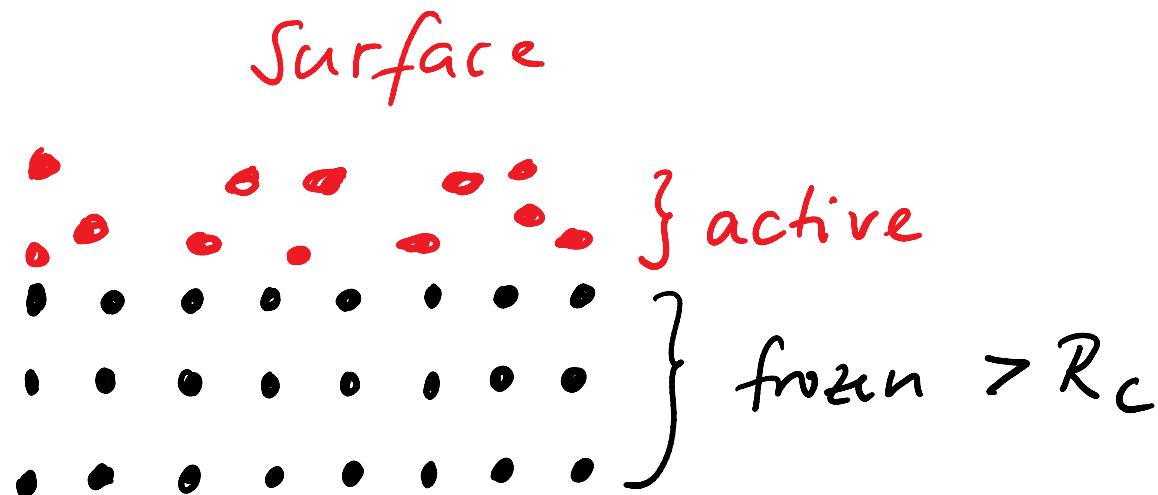
Two strategies:

1. Simulation of a slab
Periodic BC in 2 directions, open BC in remaining one

If the slab is thick enough, the inner part will look like a bulk system and we have two independent surfaces
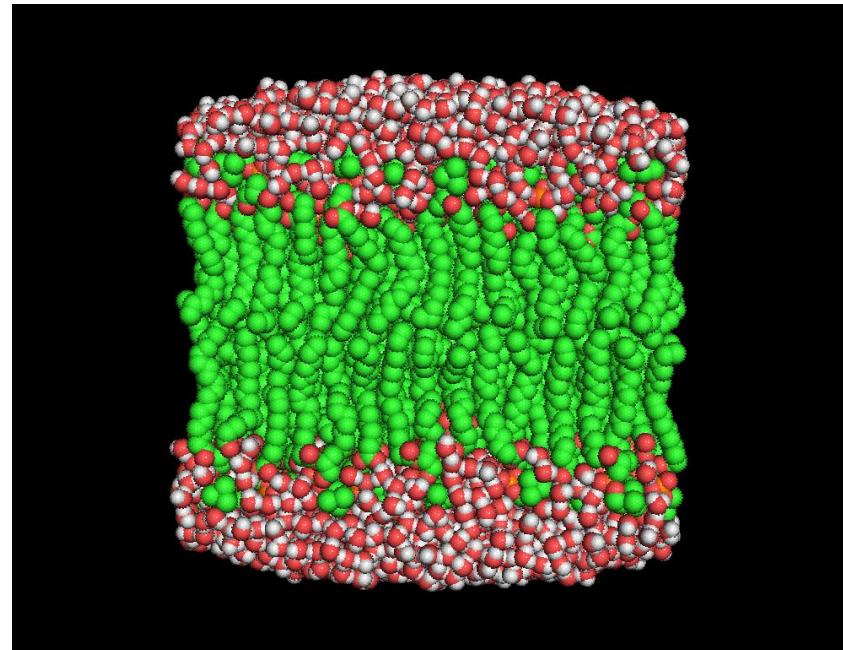
2. For static questions:

Freeze a few layers of atoms in the known bulk configuration

Surface

} active

} frozen $> R_c$

6.4 Starting and Controlling the Simulation

How to
- Initialize positions and velocities
- Equlibrate the system
- Control simulation

H Heller, M Schaefer, & K Schulten, J. Phys. Chem. 97:8343, 1993

6.4.1 Starting the simulation

Create initial set of positions and velocities:

- Positions usually defined on lattice or at random (if random avoid too short distances)
- If knowledge about the structure exists, put it in!

- Velocities are assigned random values, magnitudes reflect desired total energy or temperature
- Average (center-of-mass) velocity should be zero (otherwise you simulate translation of system as a whole)

→ This initial state is **not the equilibrium** state!
It will take the system some time to reach equilibrium.

Continuing a simulation:

Very often the best choice of initial conditions can be obtained from the previous run, if parameters have changed only slightly.

Particularly useful for large set of runs that systematically explore the parameter space.

6.4.2 Controlling the system

Thermodynamic system has a number of state variables which describe its macroscopic state such as

- Particle number, volume, temperature, pressure, total energy

They are not all independent, but connected by equations of state

Example: Ideal gas of non-interacting point particles

$$p = p(V, T)$$
$$E = E(V, T)$$

$$p \cdot V = N \cdot k_B \cdot T$$
$$p = \mathcal{S} k_B T$$
$$E = \frac{3}{2} N k_B T$$

In MD simulation: some state variables are external parameters, others are observables to be calculated

Simplest setup: **Microcanonical ensemble** (NVE)

- Particle number N
- Volume V
- Total energy E

External parameters

- Temperature T
- Pressure P

Observables to be calculated (see below)

Sometimes one wants to perform a simulation at constant T and/or constant p rather than constant E or constant V → modifications of molecular dynamics which change E and V on the go so that T and p are constant

Canonical ensemble (NVT)

- Particle number N
- Volume V
- Temperature T

External parameters

- Total energy E
- Pressure P

Observables to be calculated

Requires a **thermostat**, an algorithm that adds and removes energy to keep the temperature constant

- Velocity rescaling based on equipartition theorem

$$\left\langle \bar{E}_{kin} \right\rangle = \frac{1}{2} \sum_i m_i \left\langle \vec{v}_i^2 \right\rangle = \frac{3}{2} N k_B T$$

- Berendsen thermostat, Anderson thermostat

Isothermal–isobaric ensemble (NPT)

- Particle number N
- Pressure P
- Temperature T

External parameters

- Total energy E
- Volume V

Observables to be calculated

Requires a **barostat** in addition to the thermostat, an algorithm that changes volume to keep the pressure constant
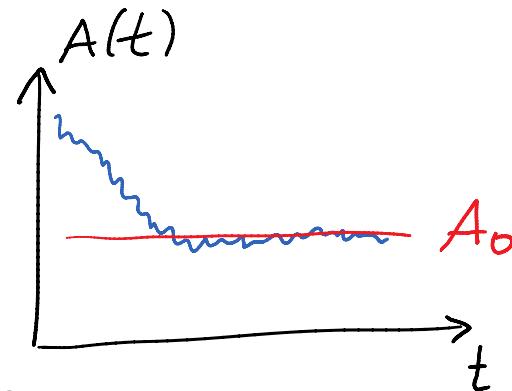
6.4.3 Equilibration

- After initial setup or after change of parameters, system is out of equilibrium. i.e. its properties will not be stationary but drift, relax towards new equilibrium state

→ if we are interested in equilibrium, must wait for a number of time steps to reach equilibrium before measuring observables

Normally, observable A(t) approaches equilibrium value A_0 as

$$A(t) = A_0 + C e^{-t/\tau_{eq}}$$

A(t) short time average to eliminate fluctuations)

What is τ_{eq} , i.e. how long do we have to wait? Hard to know a priori.

Simple estimates:

- Each particle should have had a few collisions (to exchange energy)
- Particles should have moved at least a few typical interparticle distances to explore the potential

Best solution:

- Watch a characteristic observable and monitor its approach to a constant value
- If E, N and V are fixed external parameters, you could watch T and/or p
- Compare runs with different initial conditions

6.5 Simple Observables

1. Looking at the atoms

Simplest analysis, reveals a lot about simulation
Looking at trajectories $\vec{r}_i(t)$ or $x_i(y_i), x_i(z_i)$ gives information:
- How far does atom move during run?
- Are there collisions?

Looking at configuration of all atoms at fixed time (snapshot) gives info about structure (random vs ordered…)

2. Statistical Quantities

$$A(t) = f(\vec{r}_i(t), \vec{v}_i(t))$$

- A(t) will fluctuate with t. Fluctuations are the stronger the smaller the system

- Thermodynamic behavior is represented by average:

$$\langle A \rangle = \frac{1}{N_t} \sum_{n=1}^{N_t} A(t_n) \qquad t_n = n\tau$$

- Measurement can only be started **after** equilibration

a) **Potential energy**

$$E_{pot}(t) = \frac{1}{2} \sum_{i \neq j} V(\vec{r_i} - \vec{r_j})$$

b) **Kinetic energy**

$$E_{kin}(t) = \frac{1}{2} \sum m_i \vec{v}_i^2(t)$$

c) **Total energy**

$$E(t) = E_{kin}(t) + E_{pot}(t)$$

- Should be conserved in Newton's dynamics
- Energy conservation is a good check of the time integration
- Typically relative fluctuations of, say, $\sim 10^{-4}$ are OK

$\rightarrow$ Choice of **time step** τ:

small enough to conserve energy to accuracy 10^{-4} ,

but large enough to allow for sufficiently long simulation time

$\rightarrow$ **compromise**!

d) **Temperature**:   derived quantity in MD simulation
                        in microcanonical (NVE) ensemble

Equipartition theorem: (statistical physics):
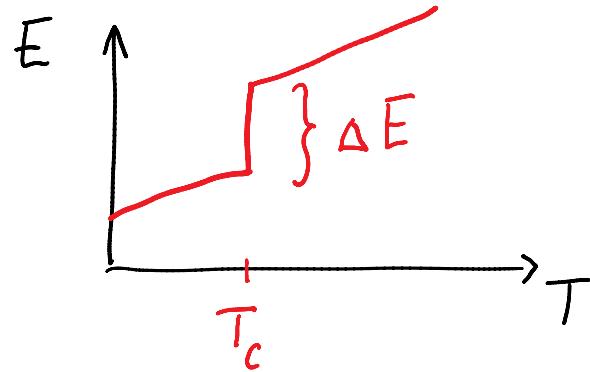Every quadratic degree of freedom takes energy ½k_BT

Kinetic energy is quadratic in v_i

$$\langle E_{kin} \rangle = \frac{3}{2} N k_B T \quad \Rightarrow \quad T = \frac{2 \langle E_{kin} \rangle}{3 N k_B}$$

e) **Caloric curve** E(T), specific heat $c_v = \frac{\partial E}{\partial T}$

E(T) will have features at a phase transition

- E(T) has jump

1st order phase transition
ΔE latent heat
Examples: melting of ice,
                    liquid water →vapor

OR:

- E(T) is continuous (no latent heat), but $\frac{\partial E}{\partial T}$ is discontinuous:
  → **continuous phase transition** (**2nd order**)
  Example: Curie point of iron

f) **Mean-square displacement**

$$\langle \Delta r^2(t) \rangle = \frac{1}{N} \sum_i \langle |\vec{r_i}(t) - \vec{r_i}(0)|^2 \rangle$$

Contains information about diffusivity, distinguishes phases

Solid: atoms remain at positions, undergo vibrations
→ $\langle \Delta r^2 \rangle$ will saturate at a value of the order of
(lattice constant)^2

Fluid: atoms can move freely
→ $\langle \Delta r^2 \rangle$ will saturate at a value of the order of
(box size)^2

Distinguish two regimes:

(i) "no collisions (small box, low density)
- Mean free path λ (distance between collisions)   **$\lambda \gg L$**
- Particles move **ballistic**, $\Delta r \sim t$
- $\langle \Delta r^2 \rangle \sim t^2$ before saturation

(ii) "many collisions (large box, high density)
- Example: real gases at ambient conditions
-          **$\lambda \ll L$**
- Particles move **diffusive**
- $\langle \Delta r^2 \rangle \sim t$ before saturation

Question: Why do we observe $\langle \Delta r^2 \rangle \sim t$ in the collision dominated regime?

- Motion can be viewed as **random walk**

$$\Delta \vec{r} = \Delta \vec{r}_1 + \Delta \vec{r}_2 + \dots + \Delta \vec{r}_N$$

$$\Delta \vec{r}_i \,\hat{=}\, \text{displacement between collisions } i-1 \text{ and } i$$

$$\Delta \vec{r}_i \,\hat{=}\, \text{random}, \quad \langle \Delta \vec{r}_i \rangle = 0, \quad \langle \Delta \vec{r}_i^2 \rangle = \Delta^2$$
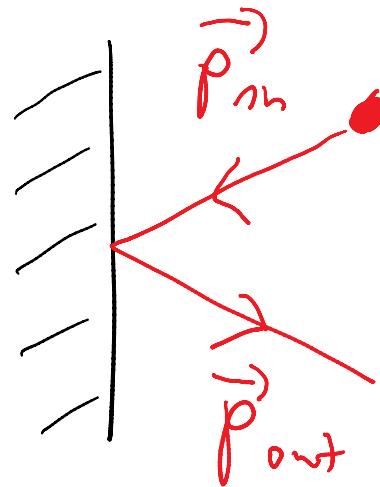
$$\Rightarrow \langle \Delta \vec{r} \rangle = 0$$

$$\Rightarrow \langle \Delta \vec{r}^2 \rangle = \langle (\Delta \vec{r}_1 + \dots \Delta \vec{r}_N)^2 \rangle = \Delta \vec{r}_1^2 + \dots + \Delta \vec{r}_N^2$$

Crossterms vanish

$$\langle \Delta \vec{r}^2 \rangle = N \Delta^2 \sim t$$

g) Pressure

Naïve idea: consider collisions of particles with walls of container, calculate force from momentum change of particles

$$\Delta \vec{p} = \vec{p}_{out} - \vec{p}_{in}$$

force on wall

$$\vec{F} = - \sum_{collisions} \Delta \vec{p} / \delta t$$

Not very efficient, only particles close to surface contribute

Better: use Clausius' virial function:

$$W\left(\vec{r}_i\right) = \sum_{i=1}^{N} \vec{r}_i \cdot \vec{F}_i^{tot} \qquad \vec{F}_i^{tot} \hat{=} \text{ total force on particle } i$$

MD average:
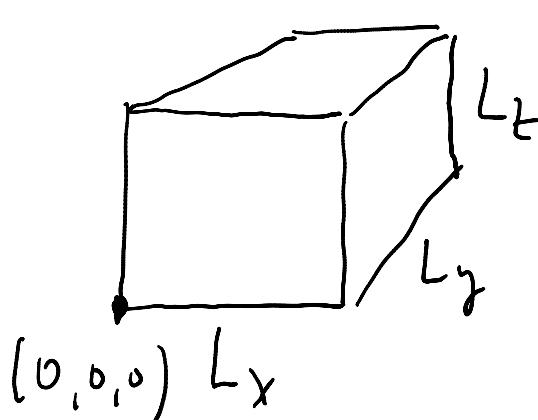
$$\langle W \rangle = \lim_{t \to \infty} \frac{1}{t} \int_0^t d\tau \sum_i \vec{r}_i(\tau) \cdot m_i \ddot{\vec{r}}_i(\tau)$$

$$= - \lim_{t \to \infty} \frac{1}{t} \int_0^t d\tau \sum_i m_i \dot{\vec{r}}_i^2 = -2 \langle E_{kin} \rangle$$

$$= - d N k_B T \qquad (\text{equipartition})$$

$$\vec{F}_i^{tot} = \vec{f}_i + \vec{f}_i^{ext}$$

$\searrow$ internal forces

$\nwarrow$ Walls!

Container: box of dimensions L_x, L_y, L_z sitting at the origin

$$\langle W^{ext} \rangle = L_x(-\rho L_y L_z) + L_y(-\rho L_x L_z) + L_t(-\rho L_x L_y) = -d\rho V$$

$(0,0,0)$ L_x

L_t

L_y

$$\langle W \rangle = \langle W^{ext} \rangle + \langle W^{int} \rangle = -d\rho V + \left\langle \sum_i \vec{r}_i \cdot \vec{F}_i \right\rangle$$

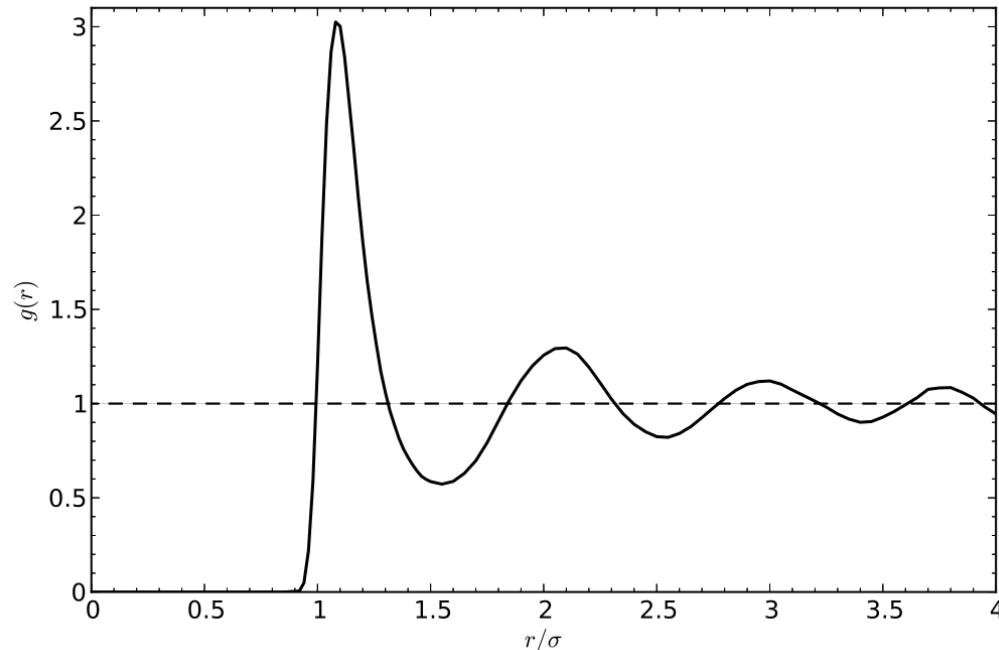$$\boxed{\rho V = N k_B T + \frac{1}{d} \left\langle \sum_i \vec{r}_i \cdot \vec{F}_i \right\rangle}$$ **Virial equation**

No internal forces ➔ ideal gas law

6.6 Real and momentum space correlations

Correlation functions:

- Relate the particle positions to each other
- Important quantities conceptually
- Directly related to scattering experiments (see chapter 5)

Pair correlation function g(r)

Describes probability for finding particle at position $\vec{r}$ if another particle is at 0 (relative to uniform distribution).

$$g(\vec{r}) = \frac{1}{\rho} \frac{1}{N} \sum_i \sum_{j \neq i} \langle \delta(\vec{r} - \vec{r}_{ij}) \rangle \qquad \rho = \frac{N}{V}$$

Particles independent, uniformly distributed:

$$\langle \delta(\vec{r} - \vec{r}_{ij}) \rangle = \frac{1}{V} \int d^3 r_{ij} \, \delta(\vec{r} - \vec{r}_{ij}) = \frac{1}{V}$$

$$g(\vec{r}) = \frac{V}{N} \frac{1}{N} \sum_i \sum_{j \neq i} \frac{1}{V} = \frac{1}{N^2} N(N-1) \longrightarrow 1$$

Any deviation from 1 describes correlations between particles!

Relation between pair correlations and structure factor

$$S(\vec{q}) = \frac{1}{N} \sum_{jk} \left\langle e^{i\vec{q}\cdot(\vec{r}_j - \vec{r}_k)} \right\rangle$$

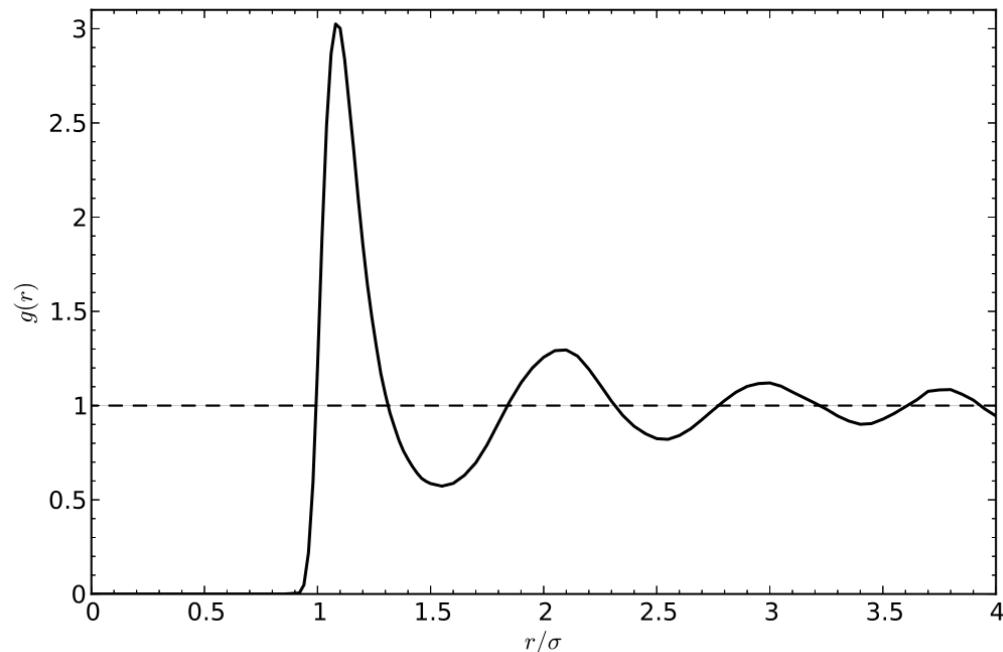Can be understood as density-density correlation function

$$S(\vec{q}) = \left\langle n(\vec{q}) \, n(-\vec{q}) \right\rangle \qquad n(\vec{q}) = \frac{1}{\sqrt{N}} \sum_j e^{-i\vec{q}\cdot\vec{r}_j}$$

Fourier transform of pair correlation function:

$$\rho \, \tilde{g}(\vec{q}) = \int d^3r \; e^{-i\vec{q}\cdot\vec{r}} \, \rho \, g(\vec{r})$$

$$= \int d^3r \; e^{-i\vec{q}\cdot\vec{r}} \frac{1}{N} \sum_j \sum_{k \neq j} \left\langle \delta(\vec{r} - (\vec{r}_j - \vec{r}_k)) \right\rangle$$

$$= \frac{1}{N} \sum_j \sum_{k \neq j} \left\langle e^{-i\vec{q}(\vec{r}_j - \vec{r}_k)} \right\rangle = S(\vec{q}) - 1$$

$$ S(\vec{q}) = 1 + \rho \, \tilde{g}(\vec{q}) $$

Measuring structure factor gives direct access to pair correlations
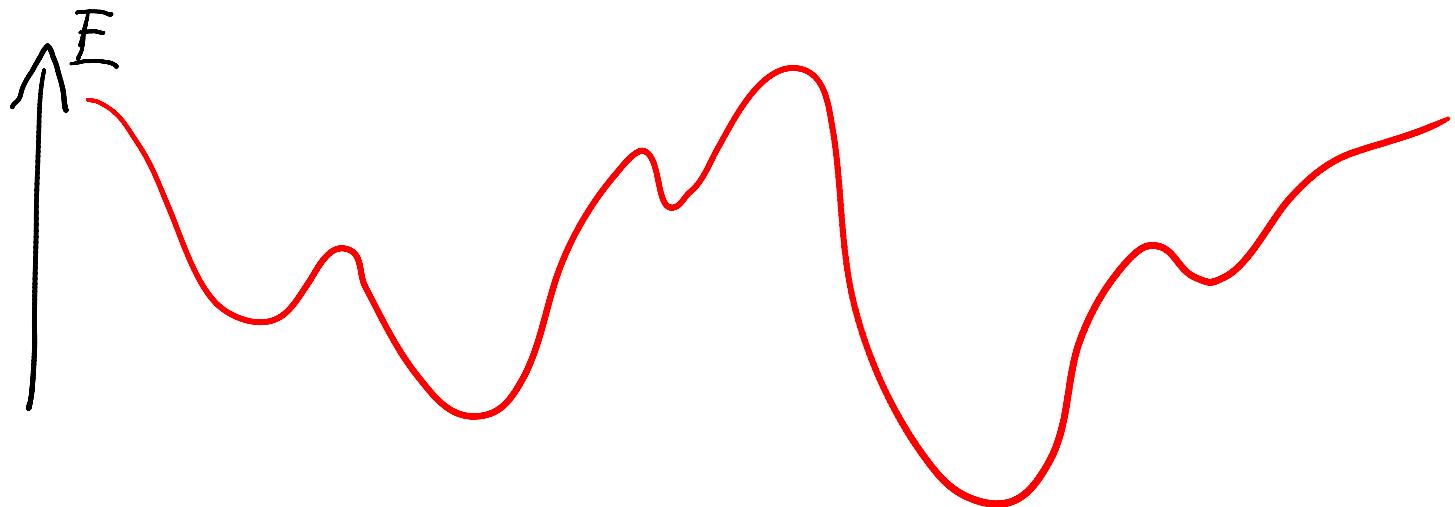
6.7 Molecular dynamics as an optimization tool

- So far, we have viewed molecular dynamics as a tool to simulate **thermodynamic equilibrium**
- Equilibration needs to be finished before measurements can begin

Now:

- Look at equilibration for its own sake
- Can be used as an **optimization algorithm**

- At **low temperatures**, equilibration means finding states with the lowest energies
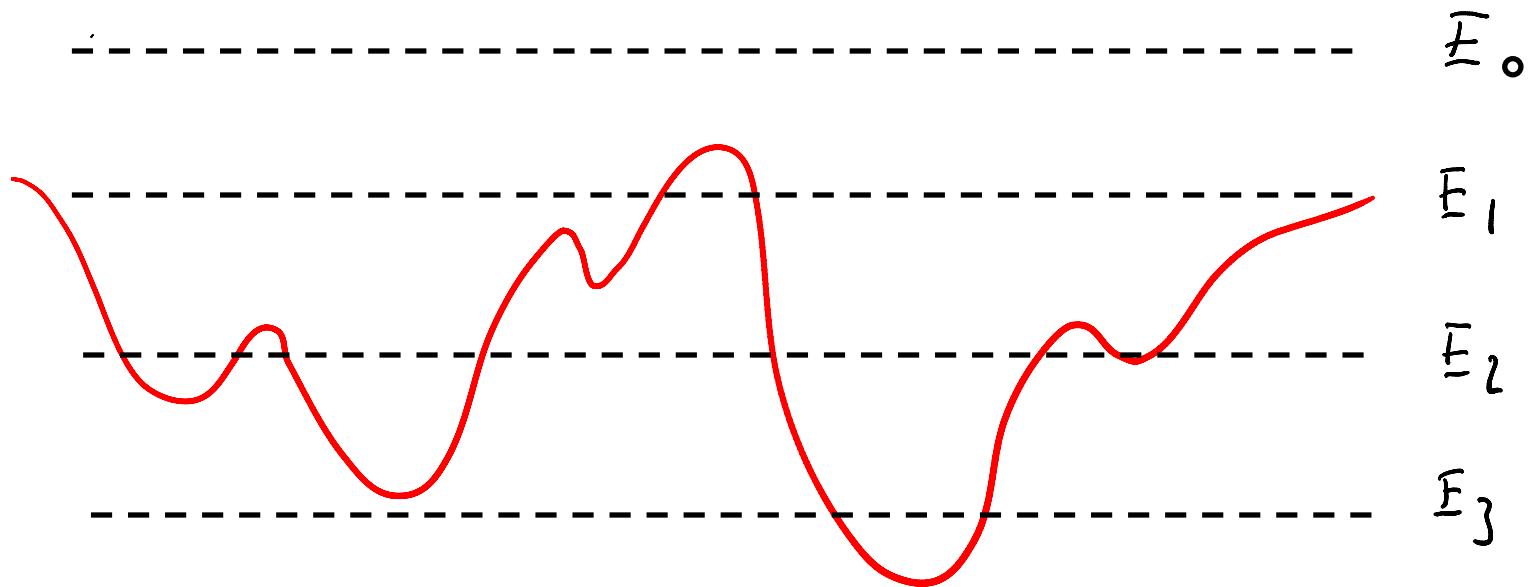- Nontrivial problem, even for small particle numbers (see project 5)

Why?

E

- Energy landscape in complicated, with many local minima
- Conventional methods (e.g., steepest descent) get stuck in side minimum

How does nature find the optimal configuration?

- System is prepared at high energy (high temperature)
- System cools down slowly = **annealing**

- At high temperatures, system easily overcomes barriers
- As temperature decreases, system will spend more time in wider and deeper minima

Simulated annealing:

- Computational algorithm that mimics annealing of a system

1. Start from a random configuration of particles and a kinetic energy larger than the typical barriers
2. Perform MD simulation, but slightly reduce kinetic energy after each time step (multiply velocities by factor a < 1)
3. Repeat until kinetic energy (temperature) is below some threshold
4. The resulting configuration is (close to) the minimum energy configuration

Remarks:

- Finding the **global** minimum is not guaranteed

- You need to **cool very slowly** to give the system time to explore the configuration space and find the deepest and broadest minima

- If you cool too quickly, system will end up in the closest local minimum rather than the global one

- velocity scaling factor a needs to be close to unity, e.g., a=0.999

Generalization: Minimum of arbitrary function F(x₁, …, x_N)

- Interpret $F(x_1, …, x_N)$ as a potential energy

- Add kinetic energy
$$E_{kin} = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{x}_i^2$$

   (masses can be chosen arbitrarily)

- Perform simulated annealing as above