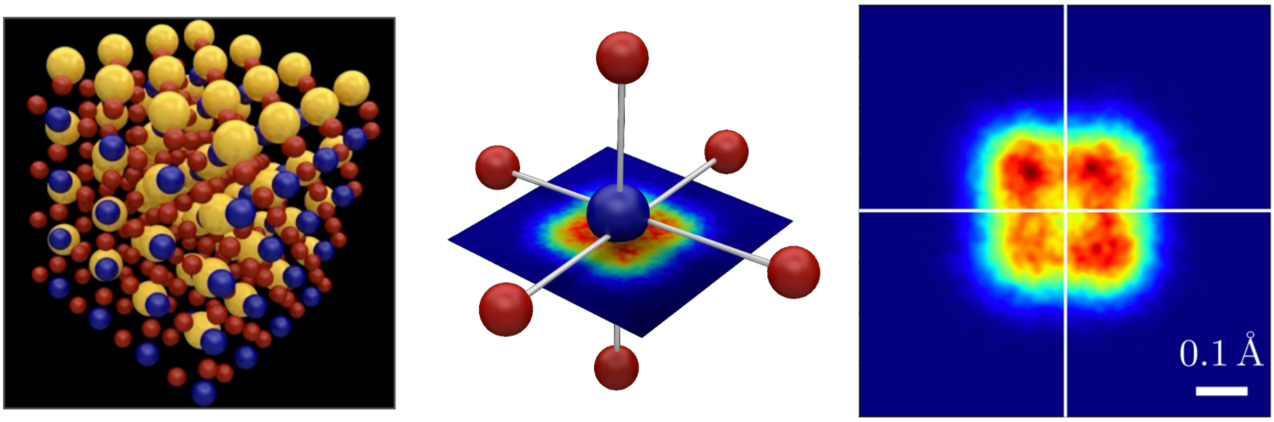
MSE 468 Lectures 12-13-14

Molecular dynamics



Molecular dynamics: Some history

- MANIAC operational at Los Alamos in 1952
- Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller (1953): Metropolis Monte Carlo method
- Alder and Wainwright (Livermore 1956): dynamics of hard spheres
- Vineyard (Brookhaven 1959-60): radiation damage in copper
- Rahman (Argonne 1964): liquid argon
- Car and Parrinello (Sissa 1985): ab-initio MD

Classical MD Bibliography

- Allen and Tildesley, *Computer Simulations of Liquids* (Oxford)
- Frenkel and Smit, *Understanding Molecular Simulations* (Academic)
- Ercolessi, *A Molecular Dynamics Primer* (http://www.fisica.uniud.it/~ercolessi/md)

Quantum MD bibliography

- Payne, Teter, Allan, Arias, and Joannopoulos, *Rev Mod Physics* 64, 1045 (1992).
- Marx and Hutter, "Ab Initio Molecular Dynamics: Theory and Implementation", in "Modern Methods and Algorithms of Quantum Chemistry" (p. 301-449), Editor: J. Grotendorst, (NIC, FZ Jülich 2000)
- http://www.theochem.ruhr-uni-bochum.de/research/marx/cprev.en.html

Decoupling ions from electrons

$$\frac{m_{\text{electrons}}}{M_{\text{nuclei}}} \approx 10^{-4}$$

- Born-Oppenheimer approximation
- We can consider (**R**$_i$: atomic configuration; ψ: electronic configuration):

$$E(\mathbf{R}_i) = \min_{\psi} E(\mathbf{R}_i, \psi)$$

(given an atomic configuration, electrons are in the ground state)

Newton's second law: *N* coupled equations

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

- Atoms treated classically (however: quantum effects can be very important for light elements, e.g. H atoms)
- The force depends on positions only (not velocities)
- The total energy of the system is conserved (microcanonical evolution)

Operational Definition

- We follow the evolution of a system composed of many **classical** particles (nuclei): [Born-Oppenheimer approximation]
- Each particle interacts simultaneously with **every other particle** (usually – but can also have 'hard spheres' contact interaction in simple models)
- Forces can be calculated with potentials or ab-initio (they are determined by electrons)
- Particles can also experience an additional external potential
- It's a **many-body** problem – albeit with a simpler informational content than in the case of the electrons (*N* particles: 3*N* coordinates, while *N* electrons: *N* wavefunctions)

Phase Space

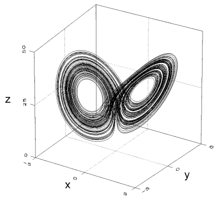
- If we have *N* particles, we need to specify positions and velocities for all of them (*6N* variables) to uniquely identify the dynamical system
- One point in a *6N* dimensional space (the phase space) represents our dynamical system

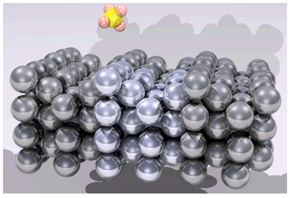
Three Main Goals

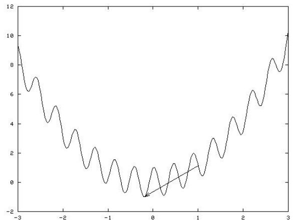
- Ensemble averages (thermodynamics) from time evolution
  Thermal expansion, diffusion, ...

- Real-time evolution (chemistry)
  Catalysis, bond breaking, ...

- Ground-state of complex structures (optimization)
  Simulated annealing







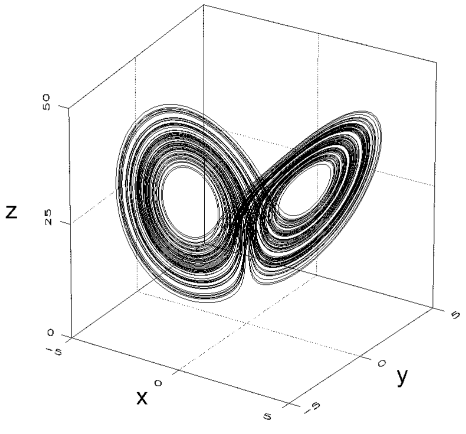
Thermodynamical averages

Under hypothesis of ergodicity, we can assume that the temporal average along a trajectory is equal to the ensemble-average over the phase space

It's like "importance sampling"

$$\langle A \rangle = \frac{\sum_i A_i e^{-\beta E_i}}{\sum_i e^{-\beta E_i}}$$

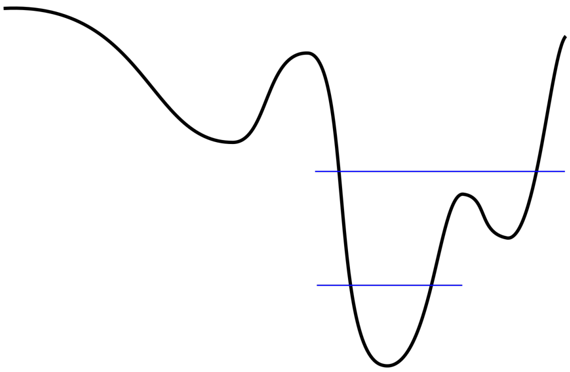
$$\bar{A} = \frac{1}{T}\int_0^T A(t)dt$$



Sampling the potential energy landscape

The choice of initial conditions and temperature determines the energy of the system

In turn, this determines the region of the potential energy landscape that will be sampled



The "Computational Experiment"

1. Initialize: select positions and velocities
2. Integrate: compute all forces, and determine new positions
3. Equilibrate: let the system reach equilibrium (i.e. lose memory of initial conditions)
4. Average: accumulate quantities of interest (thermodynamics, transport, ...)

1. Initialization

- Newton's equations: second order differential equations. Boundary conditions require initial positions and initial velocities
- Initial positions: reasonably compatible with the structure to be studied (close to minimum, possibly with small perturbations)
  - Random configurations usually not good: we need to avoid overlaps or short distances
- Initial velocities: zero, or small (according to Boltzmann distribution, for instance)
  - Can slowly heat up the system, increasing the temperature (i.e. the velocities) and thermalising until target T is reached.
  - Velocity equilibration is fast because of collisions

Maxwell-Boltzmann distribution

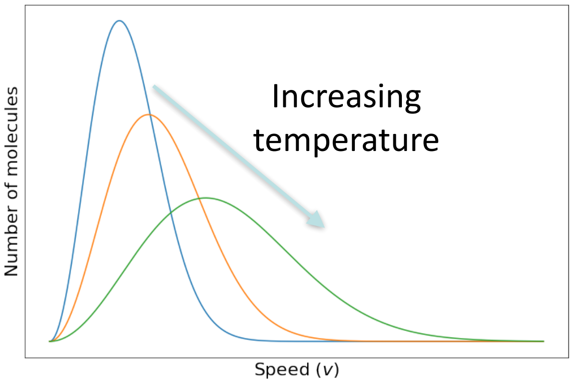
$$n(v) \propto \left(\frac{m}{2\pi k_B T}\right)^{3/2} v^2\, e^{-\frac{mv^2}{2kT}}$$

$$\bar{v} = \sqrt{\frac{2k_B T}{m}}$$

Most probable speed

$$v_{\mathrm{rms}} = \sqrt{\langle v^2 \rangle} = \sqrt{\frac{3k_B T}{m}}$$

Root mean-square speed

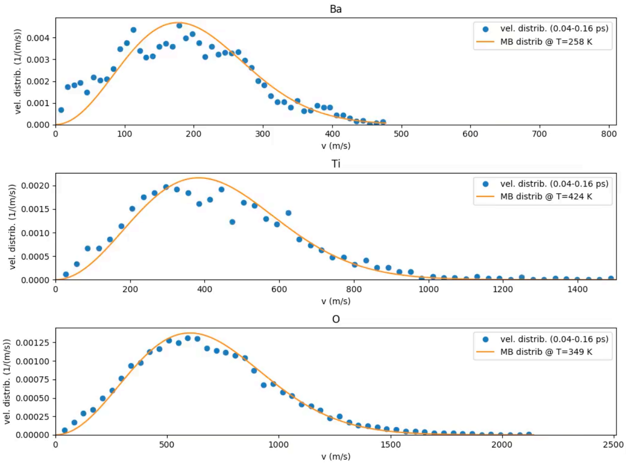


O_2 at room T: ~1500 km/h

- We can generate (with random number generator) initial velocities with norm according to this distribution
- Remove average momentum (and angular momentum) to avoid system moving (or rotating) during simulation

Maxwell-Boltzmann distribution

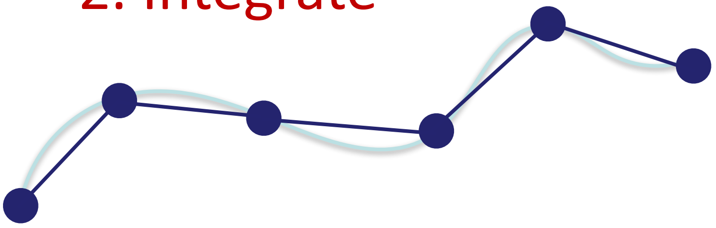
- Velocity equilibration in $BaTiO_3$
  (animation over ~7 ps of MD simulation in a 4x4x4 supercell = 320 atoms)



- Velocity equilibration is faster than position equilibration:
  - Positions evolve in a continuous way in a collision
  - Velocities could even change sign: much more randomisation

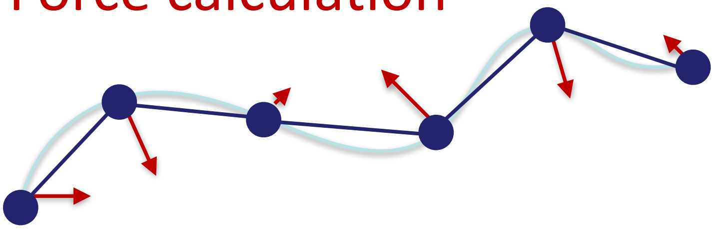
---

2. Integrate



- Time discretised in "time steps" of interval Δt
- Requirements:
  - **Time-efficient**: each step requires to compute all forces: we want a long time step
  - **Memory-efficient**: store minimal number of quantities (e.g. only the quantities at time t and t - Δt)
  - **Robust and stable**: long-term conservation of constants of motion: energy, momentum
  - **Accurate**: remain as close as possible to exact trajectory
  - **Time-reversible, constant volume in phase space**

Force calculation



- At each time step, compute the force:
  - Pair potentials (Lennard-Jones, Morse, ...)
  - Many-body potential, reactive potentials, ...
  - Tight binding
  - Quantum mechanics (e.g. DFT)

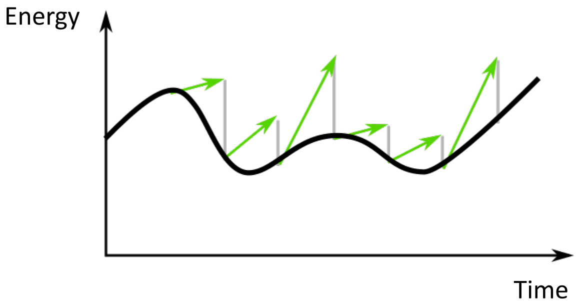
- **Note**: force calculation is the most expensive part of MD
  - Best if analytical derivatives of V are available, and if V short-range

---

Ab-initio molecular dynamics

- Born-Oppenheimer ab-initio molecular dynamics
- At each point of the trajectory, the Kohn-Sham DFT wavefuntions are *minimised to their ground state*
- Forces are computed using Hellmann-Feynman theorem

$$m_i\ddot{\mathbf{R}}_i = \mathbf{F}_i = \langle\Psi| - \frac{dV}{d\mathbf{R}_i}|\Psi\rangle$$



Integrate: discretisation of time

- Use an **integrator**: Verlet, leapfrog Verlet, velocity Verlet, Gear predictor-corrector, ...
- Choose thermodynamic ensemble: microcanonical (NVE), or canonical (NVT) using a thermostat, or isobaric-isothermic (NPT) also with a barostat…
- Types of thermostats: Stochastic (Langevin), constrained (velocity rescaling), extended system (Nose-Hoover), ...

Integrators: first-order integrator (not good)

- Start with initial conditions at *t*=0
- Recursively generate positions and velocities at successive times: from *t* to *t* + Δ*t* (Δ*t*: time step)
- Start from Taylor expansion (assumption: **r**, **v** smooth):

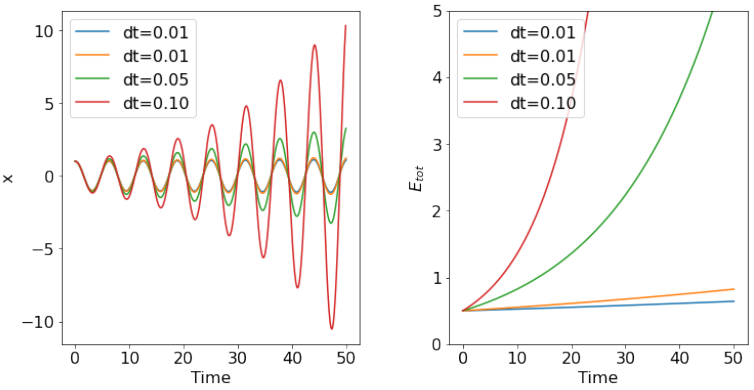
$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + O(\Delta t^2)$$

$$\mathbf{v}(t+\Delta t) = \mathbf{v}(t) + \frac{\mathbf{F}(t)}{m}\Delta t + O(\Delta t^2)$$

- This is a first-order algorithm
  - Time-step error of order $O(\Delta t^2)$

Integrators: first-order integrator (not good) Example for a Harmonic oscillator

- Example: $x(0) = 1$, $v(0) = 0$
- Exact solution should be: $x = \cos(\omega t)$, $\omega = k/m$; total energy should be conserved
- **Not very good integrator!**



Better integrator: Verlet algorithm

- Expand to higher-order, backward and forward

$\mathbf{a}(t) = -(1/m)\nabla V(\mathbf{r}(t))$

Third time-derivative of position

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + (1/2)\mathbf{a}(t)\Delta t^2 + (1/6)\mathbf{b}(t)\Delta t^3 + O(\Delta t^4)$$

$$\mathbf{r}(t-\Delta t) = \mathbf{r}(t) - \mathbf{v}(t)\Delta t + (1/2)\mathbf{a}(t)\Delta t^2 - (1/6)\mathbf{b}(t)\Delta t^3 + O(\Delta t^4)$$

- Adding the two expression gives:

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

- Odd-power terms vanish, order is **4th order**!
- Only need positions at time t and t - Δt; **no explicit use of velocities**
- I can still retrieve velocities (I need them to compute kinetic energy), *but with larger error:* $\mathbf{v}(t) = \dfrac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t-\Delta t)}{2\Delta t}$.

Variants of Verlet's Algorithms

$\mathbf{a}(t) = -(1/m)\nabla V(\mathbf{r}(t))$

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

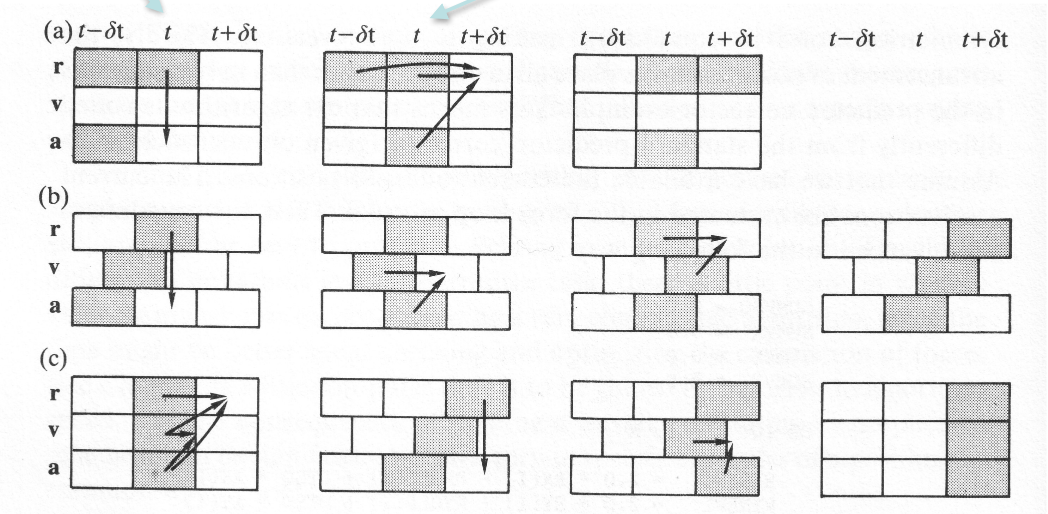


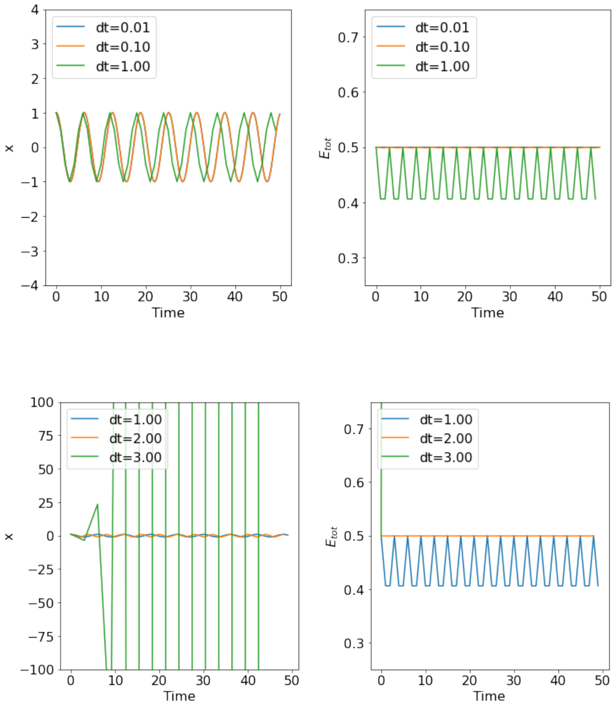
Fig. 3.2 Various forms of the Verlet algorithm. (a) Verlet's original method. (b) The leap-frog form. (c) The velocity form. We show successive steps in the implementation of each algorithm. In each case, the stored variables are in grey boxes.

- Same accuracy, same memory cost, but they allow to have also velocities

---

Verlet algorithm: Harmonic oscillator

- Allows for much longer with Δt with still quite good accuracy
- For a simple harmonic oscillator, energy is well conserved, error accumulates as phase shift
    - Still, there is accumulation of errors!
- When it fails, it fails badly! (so you notice)



Gear Predictor-Corrector

- **Predictor**: from positions and their time derivatives up to order *q* (at time *t*), predict all quantities at time $t+\Delta t$. These include the acceleration *a*.
- **Force evaluation**: compute force at predicted position. In general, will be different from predicted acceleration. The difference is an "error signal"
- **Corrector**: Use the error signal to "correct" positions and their derivatives; corrections proportional to error signal, with coefficient being a "magic number" determined to maximise algorithm stability
- It's a higher-order algorithm, but it's not time-reversible

---

Lyapunov Instabilities

- Two classical trajectories, initially very close, will eventually diverge from one another exponentially with time
- E.g. two simulations with almost identical starting conditions become statistically uncorrelated within some hundreds of steps
  - Simulated trajectory always diverges from "true" one
  - Not really an issue if we want to **collect statistics** rather than predict exact position in the future

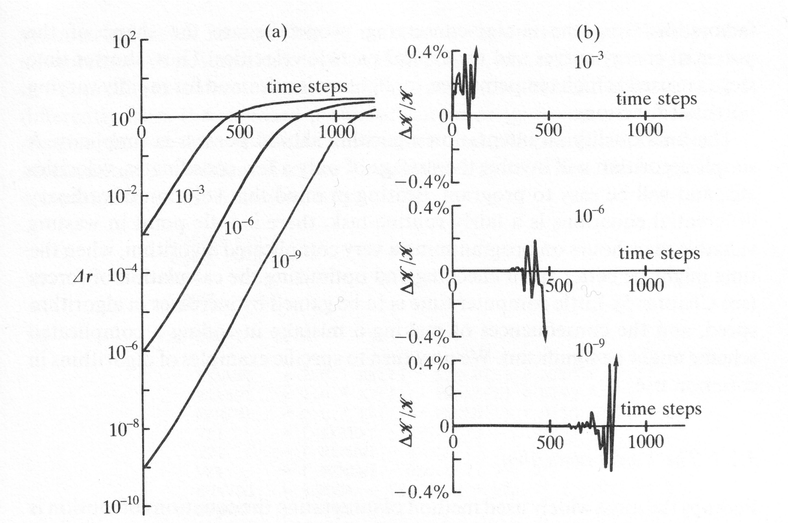


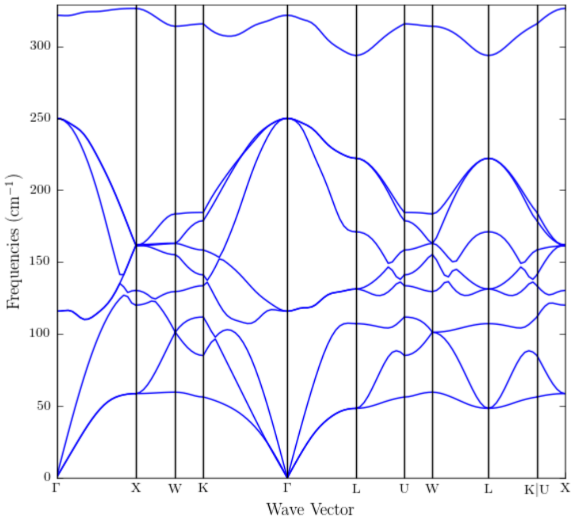
Fig. 3.1 The divergence of trajectories in molecular dynamics. Atoms interacting through the potential $v^{\mathrm{RLJ}}(r)$, eqn (1.10a), were used, and a dense fluid state was simulated ($\rho^* = 0.6$, $T^* = 1.05, \delta t^* = 0.005$). The curves are labelled with the initial displacement in units of σ. (a) Δr is the phase space separation between perturbed and reference trajectories. (b) $\Delta\mathcal{K}/\mathcal{K}$ is the percentage difference in kinetic energies.

See e.g. Allen and D. J. Tildesley Oxford University Press (2002)

Time Step

- Time step depends on the property being simulated and desired accuracy
- Too small time step: very long to simulate
- Too long time step: problems
- Typical time steps: 0.5-10 femtoseconds
    - For a liquid: small compared to mean time between collisions
    - For molecule/crystal, time step should be < 1/10 (e.g. ~1/100) of the time of the shortest period of vibrations (highest phonon frequency)
- When converging: converge at high T ("sample" higher frequencies) in small supercell (same frequencies)

Timestep and phonon frequencies



Example: **PbF_2**

ω_{max} ~ 330 cm^{-1} ~ 10 THz => Period: ~100fs

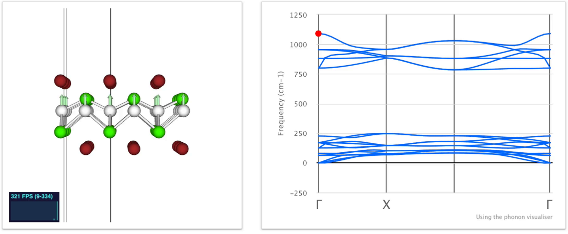
=> *use timestep (quite) smaller than 10fs*

Timestep and phonon frequencies

Ca**H**Br

$\omega_{max} \sim 1100\ \text{cm}^{-1}$

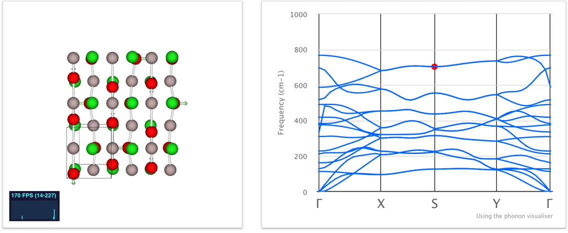
moving: H



Al**O**Cl

$\omega_{max} \sim 800\ \text{cm}^{-1}$

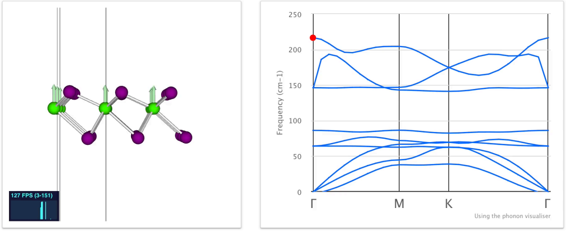
moving: O



CaI_2

$\omega_{max} \sim 220\ \text{cm}^{-1}$

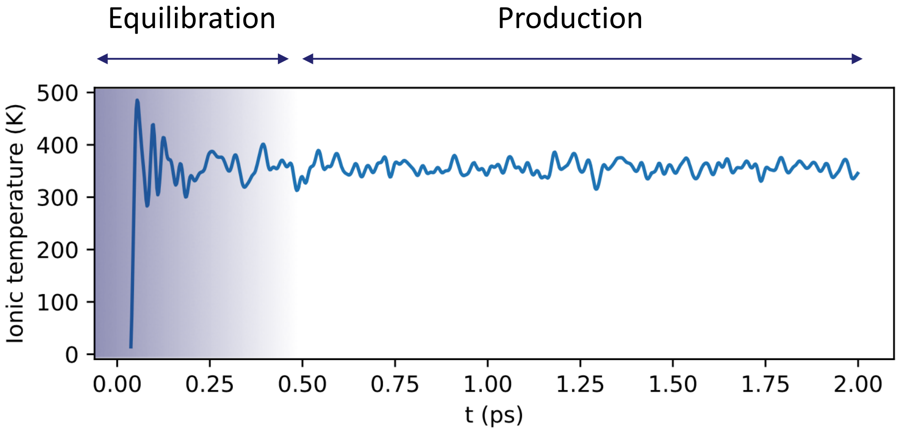
moving: Ca



- Typically the lightest mass determines the highest phonon frequency
- $\omega \sim k/m$ - note that it depends also on *k*

Phonons from Materials Cloud
https://www.materialscloud.org/discover/mc2d

3. Equilibration: How to test for it?



- Drop longer and longer initial segments of your dynamical trajectory, when accumulating averages
- Transient part of the trajectory can be long, depends on dynamics

4. Accumulate averages

- Total energy (conserved; monitor it!)
- Potential energy, kinetic energy
- Temperature (from average kinetic energy):

$$\frac{3}{2}Nk_BT = \left\langle \sum_{i=1}^{N} \frac{1}{2} m_i |\mathbf{v}_i|^2 \right\rangle$$

- Pressure (from stress)
- Caloric curve $E(T)$: latent heat of fusion
- Mean square displacements (diffusion)
- Radial (pair) distribution function

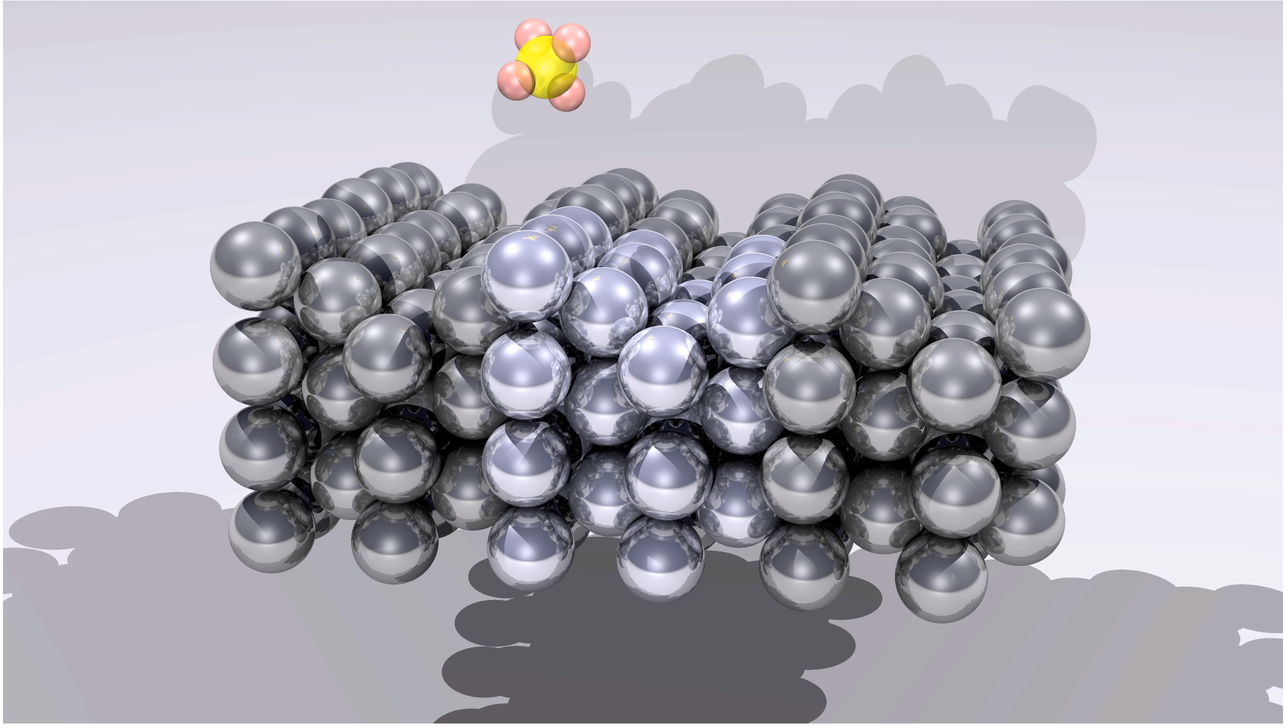
General MD workflow, correlation and restarts

1. Set particle positions and velocities
2. [start loop] Calculate forces on each particle
3. Move particles by Δt
4. (every K steps): save current positions and velocities to disk [for restarts and analysis]
5. **If** max number steps reached: go to 6, **else** go to 2
6. Stop simulation
7. Analyze results, compute averages, ...

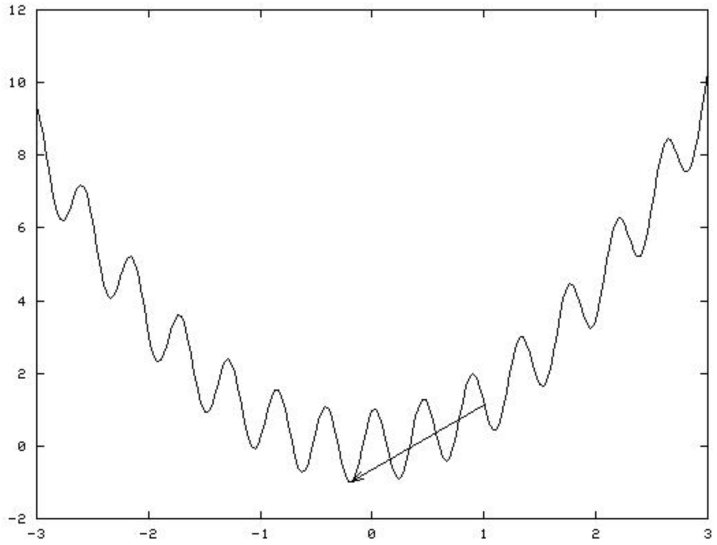
General MD workflow, correlation and restarts

- Dumping to disk can be very (space) expensive
  - No need to dump at every step (e.g.: every ~10): same information content, all quantities are **correlated**
  - Dump even more rarely for restarts (especially in DFT), e.g. every (~100)
    - Needed if e.g. you have a max running time on the queue, or if the calculation crashes
    - Reuse last-saved positions and velocities, continue dynamics (no need to equilibrate again)

Real Time Evolution



Simulated Annealing

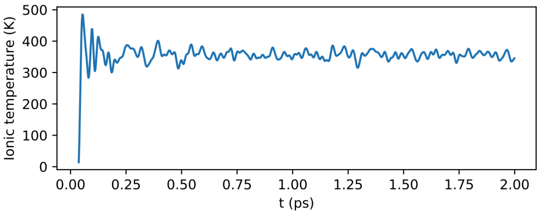


Limitations

- Time scales
- Length scales
  (PBC helps a lot, e.g. for a box of water molecules one removes surface effects)
- Accuracy of forces
- Classical nuclei

Thermostats

- When evolving the equations using Newton's law, we conserve energy: we are in a microcanonical (NVE) ensemble
- Kinetic energy (and T) will fluctuate



- Most importantly, we are sampling a different ensemble
- If we want to sample the canonical ensemble: we need to define a temperature in input, monitor it during the simulation and maintain it with a **thermostat** (feedback acting on kinetic energy of the system)

Velocity rescaling in the strong coupling

- Simple "naive" approach: since temperature is related to average kinetic energy:

$$\frac{3}{2} N k_B T = \left\langle \sum_{i=1}^{N} \frac{1}{2} m_i |\mathbf{v}_i|^2 \right\rangle$$

we can just rescale the velocities of all particles to reach the desired temperature, using a scaling factor

$$\lambda = \sqrt{\frac{T_0}{T(t)}}$$

Desired temperature

"Instantaneous" temperature

- Easy to implement, but **big disadvantage**: this strong coupling completely **spoils the natural dynamics**!

Canonical ensemble and temperature fluctuations

- Even in a canonical ensemble, the "instantaneous" temperature obtained from the equipartition theorem and the kinetic energy *will fluctuate*
  - The theorem only gives us a relation with the average kinetic energy
  - Trying to enforce that the total kinetic energy is constant is unphysical, too many constraints

$$\frac{3}{2}Nk_BT = \left\langle \sum_{i=1}^{N} \frac{1}{2}m_i|\mathbf{v}_i|^2 \right\rangle$$

- Thermodynamical definition: temperature imposed by bringing a system in contact with thermal bath; under these conditions, stats are sampled according to the canonical ensemble
- In canonical ensemble, temperature fluctuations are (C_V: specific heat):

$$\langle \Delta T^2 \rangle = \frac{k_BT^2}{NC_v}$$

- Notes: 1) there are fluctuations; 2) For infinite system, fluctuations go to zero

---

Weak coupling: Berendsen thermostat

- Weak coupling: couple weakly with heat bath, not instantaneous rescaling
- Scale velocities at each step so that:

$$\frac{dT(t)}{dt} = \frac{T_{\text{target}} - T(t)}{\tau}$$

- There is a coupling parameter τ; for τ=*dt*, it's simply a velocity scaling; tune τ carefully
  - Large τ: too weak coupling, sampling wrong ensemble
  - Small τ: too aggressive
- In all cases: energy fluctuations are strongly reduced;
  does not reproduce the canonical ensemble!
- It is however very fast; OK to use in first equilibration phase, to quickly bring the system to desired temperature

Stochastic thermostat: Andersen thermostat

- Stochastic influence on velocities: **reassign velocity of a random particle** (like a collision)
  - Probability of collisions: $P(t) = \nu e^{-\nu t}$
  - New velocity taken from Maxwell-Boltzmann distribution at desired *T*
- Between collisions: sampling a microcanonical ensemble
- Can be shown that it reproduces the canonical ensemble
- However: **do not use to compute transport properties (rates, diffusion)**!
  - Stochastic collisions create unrealistic disturbances to dynamics; e.g.: sudden random decor relation of velocities; velocity autocorrelation function decays too fast, ...

Nosé-Hoover thermostat

- Possible to rigorously maintain the canonical ensemble: **extended system method**
- Introduced by Nosé, further developed by Hoover
- **Idea**: add explicitly a "heat bath": an additional degree of freedom to the system
  - Represents a friction term, with a mass and coupled to the system, that slows down or accelerates the system to reach target value of *T*
  - Full system: evolves microcanonically
  - *Energy exchanged between real system and heat bath*
  - Real system samples canonical ensemble

Lagrangian formalism - reminder

- Construct the Lagrangian: $L = T(q, \dot{q}) - V(q, \dot{q})$
- Equations of motion are given by:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0$$

- Convenient to use generalised coordinates
- One just needs to think at two scalar coordinates *T* and *V*

Lagrangian formalism - reminder

- Example in 1D: $T = \frac{1}{2}mv^2 = \frac{1}{2}m\dot{x}^2; \quad V = V(x)$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0$$

$$\frac{\partial L}{\partial \dot{x}} = \frac{\partial\left(\frac{1}{2}m\dot{x}^2\right)}{\partial \dot{x}} = m\dot{x} \quad \Rightarrow \quad \boxed{\frac{d}{dt}(m\dot{x}) = -\frac{\partial V}{\partial x}}$$

Newton's 2nd law

Lagrangian formalism - reminder

- One can derive Hamiltonian dynamics via Legendre's transform
- Twice as many differential equations, but all first order

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \qquad H(q,p,t) = \sum_i \dot{q}_i p_i - L(q,\dot{q},t)$$

$$\boxed{\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}}$$

- In the 1D example of before:

$$p = m\dot{x} (= mv); \quad H = \frac{p^2}{2m} + V(x) \quad \Rightarrow \quad \dot{x} = \frac{p}{m}, \quad \dot{p} = -\frac{\partial V}{\partial x}$$

Nosé-Hoover thermostat

- Nosé Lagrangian:

$$L_{\text{Nosé}} = \sum_i \frac{1}{2} m_i s^2 \dot{r}_i^2 - V(r_1, \cdots, r_n) + \frac{1}{2} Q \dot{s}^2 - \frac{g}{\beta} \ln s$$

- *Q*: plays the role of an "effective" mass associated with the new degree of freedom (*g*: number of degrees of freedom, e.g. 3*N*-3 when fixing the center of mass)
- Hoover: rewrite in terms of a "friction term"

$$\xi = \dot{s}/s = d(\ln s)/dt$$

(indeed, one of equations we get is: $\dot{p}_i = -\frac{\partial V}{\partial r_i} - \xi p_i$)

Nosé-Hoover thermostat

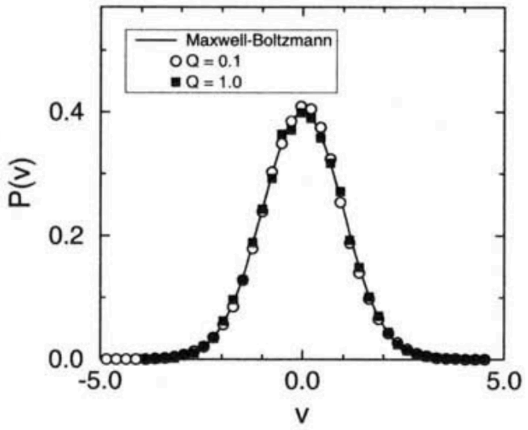
- We get also the equations of motion for ξ
- There is a conserved quantity (energy-like, but it's *not* an Hamiltonian: do *not* derive equations of motion from it; here, $p_s = Q\xi$):

$$\sum_i^N \frac{p_i^2}{2m_i} + V(r_1, \cdots, r_n) + \frac{p_s^2}{2Q} + g\frac{\ln s}{\beta}$$

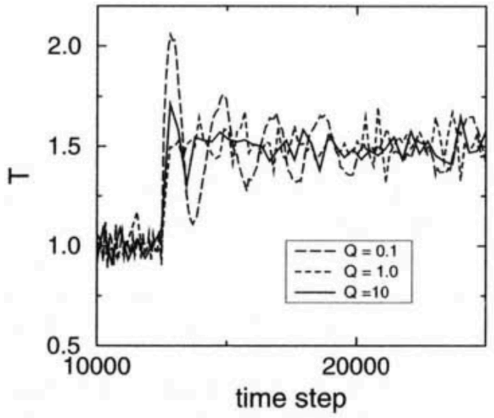
- Full system has 6N+2 degrees of freedom and evolves microcanonically; one can show that our system describes a *canonical* ensemble instead, thanks to the (ln *s*) term

---

Nosé-Hoover thermostat



Velocity distribution for a LJ fluid, vs the Maxwell-Boltzmann distribution, resulting from a simulation using a Nosé-Hoover thermostat: **results are independent of the coupling *Q***



Response of the system to a sudden increase of imposed temperature (from 1 to 1.5) for a LJ fluid over time, for various values of the Nosé-Hoover coupling *Q*.

- Small *Q*: low inertia of heat bath, long-lived/weakly-damped T fluctuations
- Large *Q*: we regain conventional MD

From Frenkel and Smit, *Understanding Molecular Simulations* (Academic)

Nosé-Hoover thermostat

- Final notes
  - It can be non ergodic or at least very slow in sampling the full equilibrium distribution (e.g. it does not work for a single harmonic oscillator, or small and stiff molecules)
  - Extensions: **Nosé-Hoover chains**: couple first degree of freedom to a second one, and repeat; resulting dynamics is sufficiently chaotic to improve sampling

---

Effect of thermostats on Lennard-Jones fluid

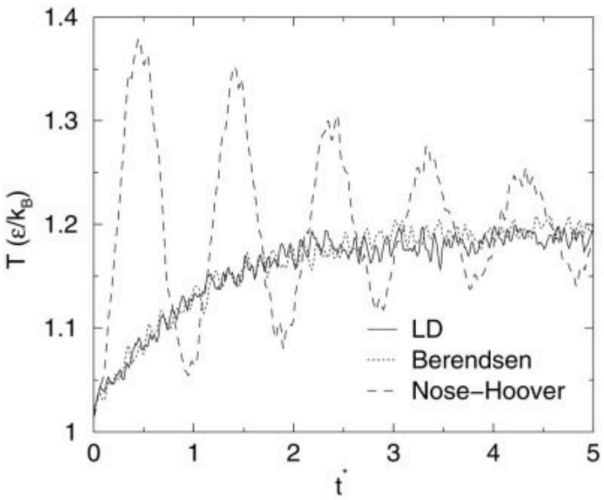


Figure 2. Response of three different thermostats to a temperature jump from 1 to 1.2 for a system of 8000 Lennard–Jones particles. Time and temperature are in reduced units (from Hess[61]).

From D. van der Spoel *et al.*, J. Comp. Chem 26, 1701 (2005)

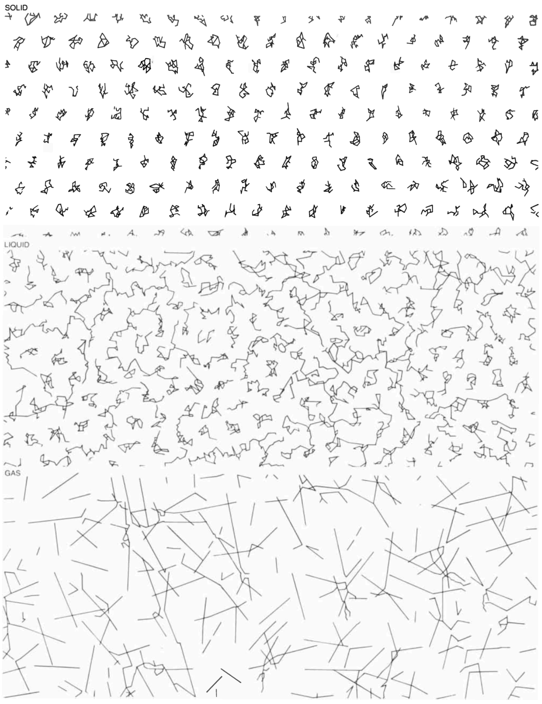
(LD: Langevin dynamics)

Parrinello-Rahman Barostat

- If you want to do a NPT simulation, you also need a barostat
- Similar idea of Nosé-Hoover can be used also to control pressure
- Parrinello-Rahman barostat, J. Appl. Phys. 52, 7182 (1981): each unit vector of the periodic cell is independent
  - Volume can vary (to keep pressure constant)
  - But also: cell shape can change (full control over stress tensor, rather than just the pressure)

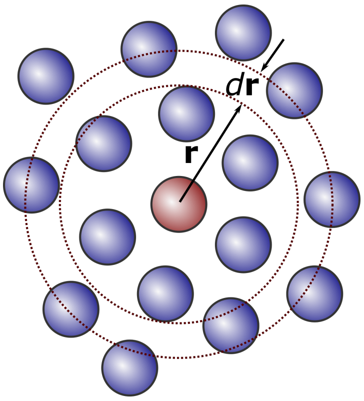
Phases of matter

- **Solids (crystals)**: Regular ordered structure; small vibrations around the equilibrium lattice position
- **Liquids**: Particles follow Brownian motion (many collisions), diffusive process
- **Gas**: Very long free paths, interrupted by relatively rare collisions



J. A. Barker, D. Henderson, Sci. American 245, 130 (1981)

Correlation functions: Radial distribution function (RDF)



- Number *dn*(*r*) of particles between *r* and *r+dr* increases with *r* (volume of the shell: $4\pi r^2 dr$)
- RDF *g*(r): is a **pair correlation function**
- RDF is a distribution function measuring the probability of finding two particles at given distance *r*, normalised with respect to system average density *N*/*V*:

Number of particles between *r* and *r+dr*

$$\frac{dn(r)}{4\pi r^2 dr}\,\frac{1}{N/V} = g(r) \qquad \Rightarrow dn(r) = \frac{N}{V} g(r) 4\pi r^2 dr$$

Volume of the shell

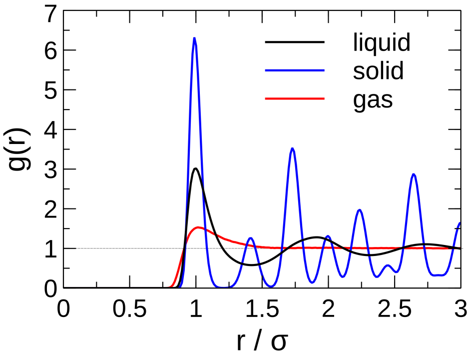
Average density

RDF

Can be used to compute coordination number:

RDF measures the local structure

- By construction: *g*(*r*) is 1 in an ideal homogeneous uniform gas with the same density
- Deviations from 1 indicate correlations between particles due to intermolecular interactions
- *g*(*r*) tends to 1 for large *r*; *g*(*r*) goes to zero for small *r* (atoms don't overlap)
- Sharp peaks for **crystalline solid** (first neighbour shell, second neighbours, ...); shallow peaks for **liquid** (there is still some structure); smooth curve in a **gas**

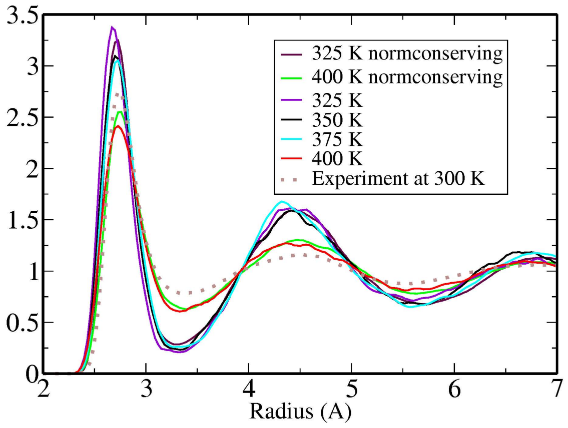


$$\frac{dn(r)}{4\pi r^2 dr}\,\frac{1}{N/V} = g(r) \qquad \Rightarrow 4\pi\rho\int g(r) r^2 dr$$

can be used to estimate **coordination number** (e.g. integrating from 0 to first minimum)

Water with ab-initio MD

- Can check the cell size (in PBC) to simulate an infinite sample
- There is a phase transition (but temperature is off in PBE)
- Nuclear quantum effects are important



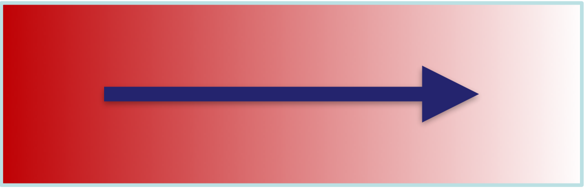
Non-equilibrium properties

- We used thermodynamic equilibrium in MD
- How can we characterise non-equilibrium processes?
- Examples of kinetic properties:
  - Changes of concentration: (bulk, surface, ...) diffusion
  - Changes of species: reaction rates
  - Changes of structure

Diffusion

Diffusion

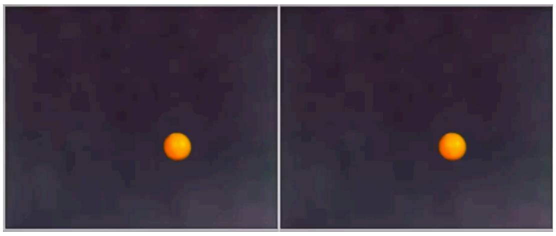
- Particles move from a high-concentration to a low-concentration region
  - **Macroscopically**: measure diffusion by change in concentration
  - **Microscopically**: diffusion is a spontaneous net movement of particles



High concentration

Low concentration

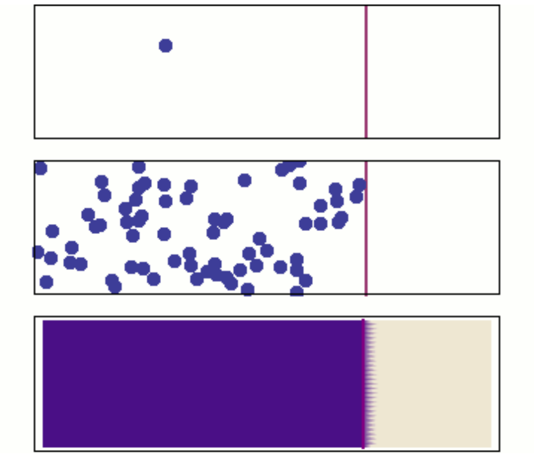
- Result of **Brownian motion**: random motion of particles



Video from Wikipedia

Microscopic and macroscopic view

- Particles on the left only, barrier removed at certain time
- Statistical trend: fill uniformly the container
- Few particles: random motion. Many molecules: randomness becomes a smooth and deterministic movement from high- to low-concentration regions
- **No microscopic force** pushing molecules to the right, but macroscopically it "seems" there is one (entropy)

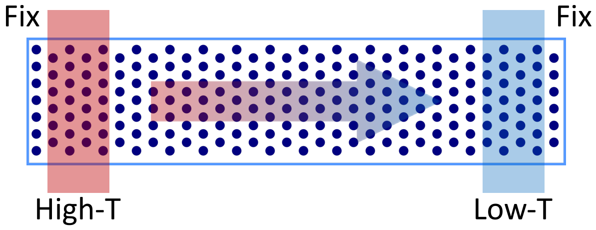


Video from Wikipedia

How to simulate

- **Method 1**: establish and maintain non-equilibrium steady state
  - Non-equilibrium molecular dynamics
  - Must continuously add and remove conserved quantities, only 1 property measured at given time, good signal-to-noise ratio (with large perturbations)

Apply thermostats to different parts of the system, maintain temperature gradient



- **Method 2**: Rely on natural fluctuations
  - All systems have fluctuations of mass, momentum, energy, ...
  - Observe statistics of these natural fluctuations; they are related to transport coefficients

Diffusion from random walk

- **Microscopic** interpretation:
  - Random walk: Particles move freely for a given length before being randomly scattered
- Fick's laws describe the **macroscopic** diffusion:

$$\frac{\partial c(\mathbf{r},t)}{\partial t} = D\nabla^2 c(\mathbf{r},t)$$

  - *c*(**r**, *t*) can be any scalar extensive quantity: number of particles, mass, energy, electric charge, ...
  - *D*: diffusion coefficient

Macroscopic diffusion

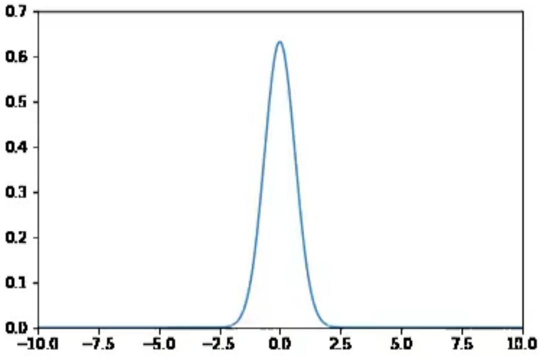
$$\frac{\partial c(\mathbf{r},t)}{\partial t} = D\nabla^2 c(\mathbf{r},t)$$

- Boundary conditions: concentration = 0 at infinity
- Solution: Gaussian spreading in time (independent of the details of the collisions, as long as they are random)

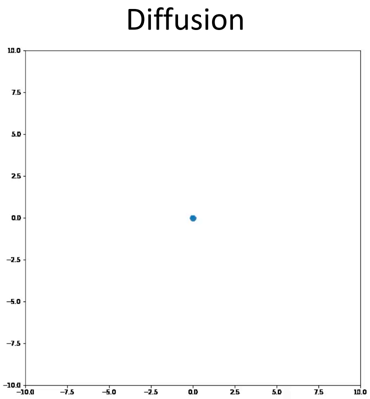
Solution in 1D

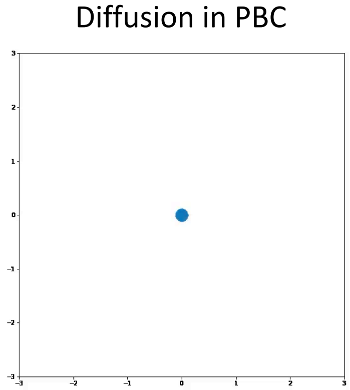
$$c(x,t) = \frac{N}{\sqrt{2\pi\sigma^2(t)}} e^{-x^2/2\sigma^2(t)}$$

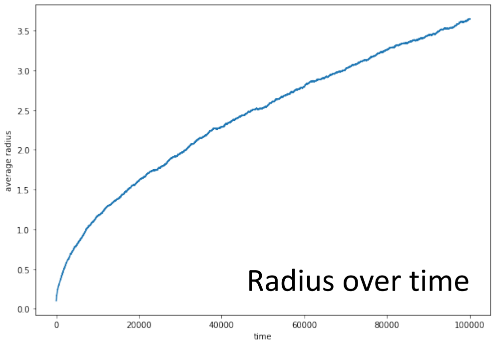
$$\sigma^2(t) = 2Dt$$



Microscopic: 2D random walk and diffusion







- Brownian motion leads to particle net motion

$$R \propto \sqrt{t}$$

Relation between microscopic and macroscopic quantities

- Fick's law:

$$\frac{\partial c(\mathbf{r},t)}{\partial t} = D\nabla^2 c(\mathbf{r},t)$$

- Multiply by r^2 and integrate over all space:

$$\frac{\partial}{\partial t}\int r^2 c(\mathbf{r},t)d\mathbf{r} = D\int r^2\nabla^2 c(\mathbf{r},t)d\mathbf{r}$$

- Integrate by parts twice on the right (*d*: dimensionality):

$$\frac{\partial}{\partial t}\langle r^2(t)\rangle = 2dD$$

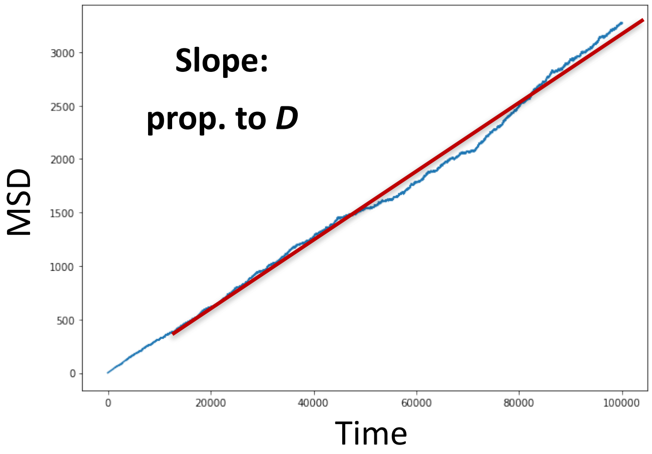
Mean square displacement

(Operative) definition of the Mean Square Displacement (MSD)

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

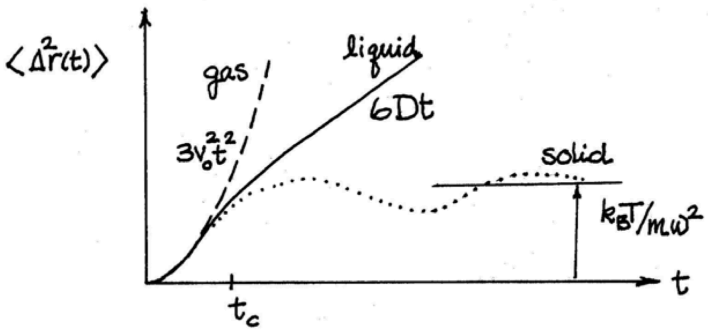
- Use MD to simulate motions of atoms
- Follow trajectories, monitor how atoms move away from their initial position *on average*
- Obtain diffusion coefficient

$$D = \frac{1}{2d} \lim_{t \to \infty} \frac{\partial}{\partial t} \langle \Delta r^2 \rangle$$

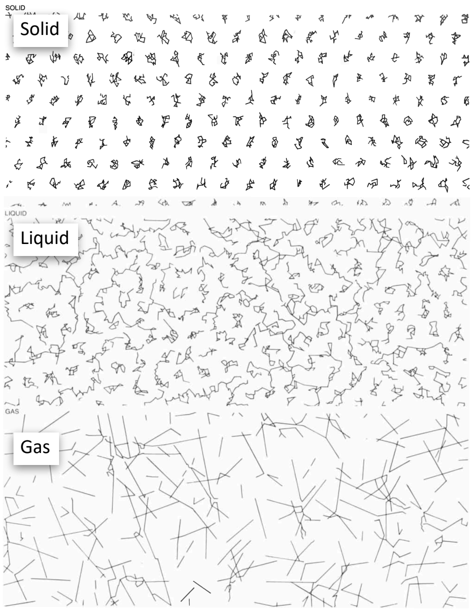


Slope:
prop. to *D*

Mean square displacement in various phases



From S. Yip's lecture notes:
https://ocw.mit.edu/courses/nuclear-engineering/22-103-microscopic-theory-of-transport-fall-2003/lecture-notes/lec2.pdf



J. A. Barker, D. Henderson, Sci. American 245, 130 (1981)

Measuring the melting temperature and nucleation issues

- Temperature at which MSD becomes linear, caloric curve has a jump, ...: typically higher in MD than real melting temperature T_m
- Melting (def): temperature at which solid and liquid coexist
- However: without a **seed** (for the solid or the liquid to grow): we see overheating above melting (metastable phase, but can exist on the time of the simulation), up to 20-30% of T_m

---

Measuring the melting temperature and nucleation issues

- Similar effect when cooling down to melt
- One approach to obtain more accurate T_m: simulate a system 50% solid, 50% liquid
- Establish equilibrium where neither of the two grows with simulation time: it's T_m
- Note: volume might change at phase transition: better to do the procedure at constant pressure

Velocity autocorrelation function (VAF)

- Let's write the displacement as an integral of the velocity

$$\Delta x(t) \equiv x(t) - x(t=0) = \int_0^t dt' v_x(t')$$

- Therefore:

$$\langle \Delta x(t)^2 \rangle = \left\langle \left(\int_0^t dt' v_x(t') \right)^2 \right\rangle = \int_0^t dt' \int_0^t dt'' \langle v_x(t') v_x(t'') \rangle =$$

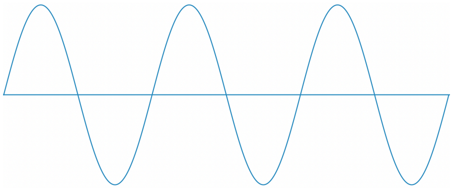
$$= 2 \int_0^t dt' \int_0^{t'} dt'' \langle v_x(t') v_x(t'') \rangle$$

- Using now the relation with the diffusion coefficient (in 1D here):

$$2D = \lim_{t\to\infty} \frac{\partial \langle \Delta x(t)^2 \rangle}{\partial t} = \lim_{t\to\infty} 2 \int_0^t dt' \langle v_x(t) v_x(t') \rangle$$

Velocity autocorrelation function (VAF)

- $\boxed{\langle v_x(t) v_x(t') \rangle}$ indicates correlation of the velocities at time *t* and *t'*; ***auto***correlation because it's the correlation of the velocity with itself
- It does not depend on the origin of time, so $\langle v_x(t) v_x(t') \rangle = \langle v_x(t-t') v_x(0) \rangle$
- E.g.: *t'*=*t*: perfect correlation. Normalising with $\langle v_x(0) v_x(0) \rangle$, the VAF is 1
- If I have periodic motion with period T, then:
  - $\frac{\langle v_x(T) v_x(0) \rangle}{\langle v_x(0) v_x(0) \rangle} = 1$ (perfectly correlated after 1 exact period)
  - $\frac{\langle v_x(T/2) v_x(0) \rangle}{\langle v_x(0) v_x(0) \rangle} = -1$ (perfectly anticorrelated after half period)
  - …



Velocity autocorrelation function (VAF)

- Calling τ=*t*-*t*' and changing variables we get:

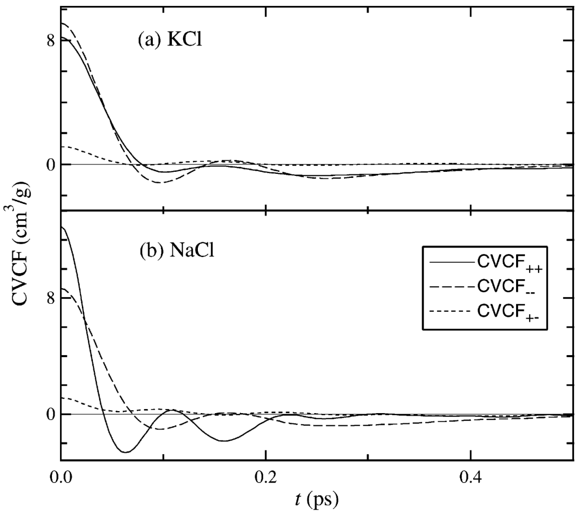
$$D = \int_0^\infty d\tau \langle v_x(\tau) v_x(0) \rangle$$

- We can get a diffusion coefficient from the autocorrelation of velocities!
- NOTE: Time correlation functions in equilibrium MD have large statistical error: they represent average response to naturally occurring fluctuations, which are small

---

Velocity autocorrelation function (VAF)

- VAF: ~ps scale (typical MD step: ~fs); integral gives us *D*
- Measure of **correlation time** of the simulation: if we measure velocity very often (timestep of ~fs), the velocities are correlated between them with a timescale of the VAF
- **=> To do good averaging**: we need simulations that are *much longer than the typical correlation times*



D. Wheeler, *et al. J. Phys. Chem. B* **108**, 18353-18361 (2004)

Green-Kubo relations

- This is one example of the more general Green-Kubo relations: exact mathematical expressions for transport coefficients in terms of integrals of time correlation functions

$$L = \beta V \int_0^\infty dt \langle J(t) J(0) \rangle$$

- This is a case of the “fluctuation-dissipation theorem”
  - Fluctuation of particle at rest has same origin as dissipative frictional force: same random forces that cause the erratic motion (*fluctuations*), e.g. Brownian motion, also cause drag (friction, *dissipation*) if particle is pulled through the fluid

---

More Green-Kubo relations

- Other transport coefficients can also be obtained via Green-Kubo relations:
  - Shear viscosity, from the stress
  - Electrical conductivity, from the charge current
  - IR adsorption, from the polarization
  - ...

Arrhenius equation for the diffusion coefficient

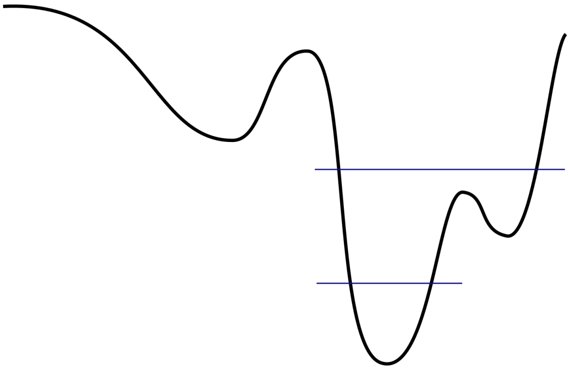
- Diffusion rates very much dependent on temperature

$$D(T) = D_0 e^{-\frac{E_A}{k_B T}}$$

- E_A: activation energy for the reaction
- To measure (D_0 and) E_A: log-linear plot as a function of $1/T$ (or most commonly $1000/T$ in units of kelvin)
    - Think at why this helps extracting, e.g., E_A!
- **If** system satisfies Arrhenius equation, can be used to "accelerate" simulations, e.g. ageing of Li batteries
    - Simulate at higher T (for many T), where shorter simulations are enough; then extrapolate to lower T
    - *Be careful if Arrhenius model works for your system!*

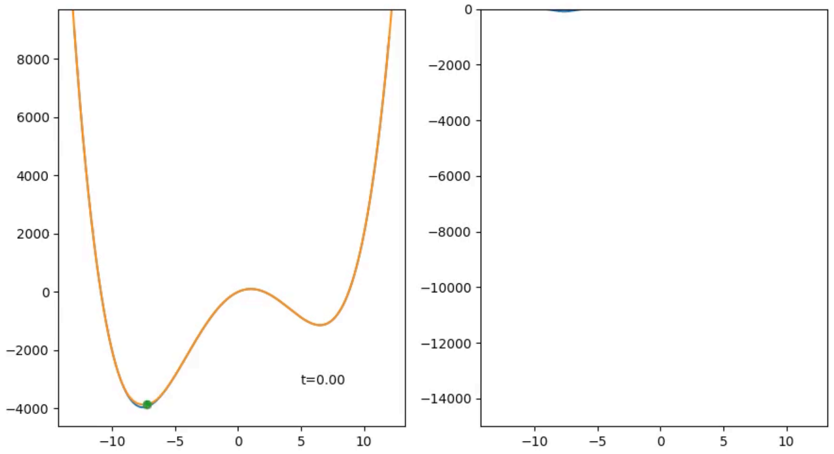
Accelerating dynamics for rare events

- If you have a high barrier, you hop/tunnel very rarely, possibly never in a MD run
- Need a way to facilitate transitions to perform a more accurate sampling of the energy landscape



Meta-dynamics

- Idea: filling the free energy wells with "computational sand" (Gaussians)
- Enhance the system sampling by discouraging revisiting of sampled states



Meta-dynamics and collective variables

- In 1D: easy
- In many dimensions: need to define "collective variables" *s* (few variables that well represent the relevant degrees of freedom, e.g. rotation of a molecule, bond length, ...)
- $H = T + V + V_{\text{bias}}(s)$; at infinitely long simulation time the accumulated bias potential converges to free energy with opposite sign
  - Laio, Parrinello, PNAS 99, 12562 (2002)
- **Most difficult part of metadynamics: selection of "correct" collective variables**
- Note: Biasing the dynamics helps explore the energy landscape and sample the free energy, but we lose dynamics, i.e. time scale
  - Transition rates are artificially accelerated, time is not uniform