

Welcome to CHE 384T: Computational Methods in Materials Science

Molecular Dynamics

LeSar Ch. 6; App. I4, G



The University of Texas at Austin
McKetta Department
of Chemical Engineering
Cockrell School of Engineering

Announcements

Some common observations for the Homework reports:

- Present polished figures (e.g., appropriate font sizes for readability, legend labels)
- Present only figures that are the final version
- Treat the report as a narrative
- Interpretation of results counts towards part of the grade
- Report should be stand-alone (i.e., one does not need the assignment to understand the report)
- If you include code as part of your report, describe its basic function
- Be sure to comment your code
- If you deviate from the provided code, describe what you did (e.g., psuedocode)
- Assume the reader/grader has little prior knowledge of the homework assignment

Goal: Improve readability

Announcements

Time to start thinking about your final project:

- Work in groups of 1-2 people
- Projects should reflect about a months worth of course time effort and should be proportional to your initial expertise (e.g., if you come from a DFT group, your final project should not be just running DFT calculations)
- Groups of 2 should propose a project that is commensurate in work with having two people on the same project. The division of work should be made clear in carrying out the project, the report, and the presentation
- Can be related to your research

Due Nov 4, 11:59pm: project proposal; [Proposal guidelines](#)

[Final Report Rubric](#)

[Final Project Presentation](#) (~12 min per presentation; 10 min presentation, 2 min Q&A)

10/25

* Mid-semester survey

* OH; CPE 4.450

Wed. 10/23

11am

11/13 10am

10/30

12pm

11/20 10am

11/16

10am

12/4 10am

Example project topics:

- Water-splitting surface reactions
- Comparison of crystalline and amorphous structure
- Grain Growth in polycrystalline structures
- Magnetic state phase diagrams with doping

Lecture Outline

What is temperature

Methods without an explicit T

Basics: Equations of Motion

Newton's equations

Verlet algorithm and other integrators

Conservation Laws

An example with Lennard-Jonesium

Initialization- positions, velocities

Cutoffs

Equilibration

Statistical analysis:

averaging, variations

ergodicity, fluctuations

Velocity and structural correlations

Basics: Thermodynamics

Laws of thermo

Entropy

Some stat mech

Thermodynamic ensembles: NVE, NVT, NPT

Thermostats:

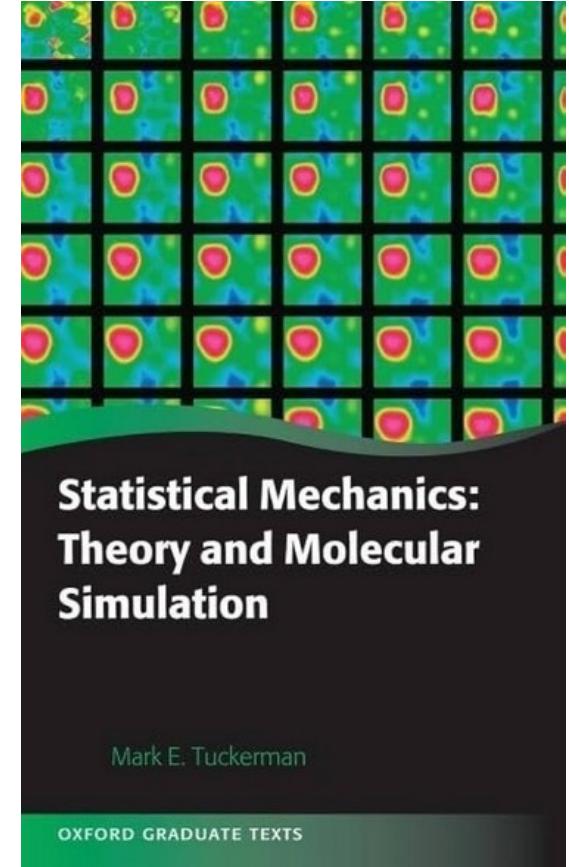
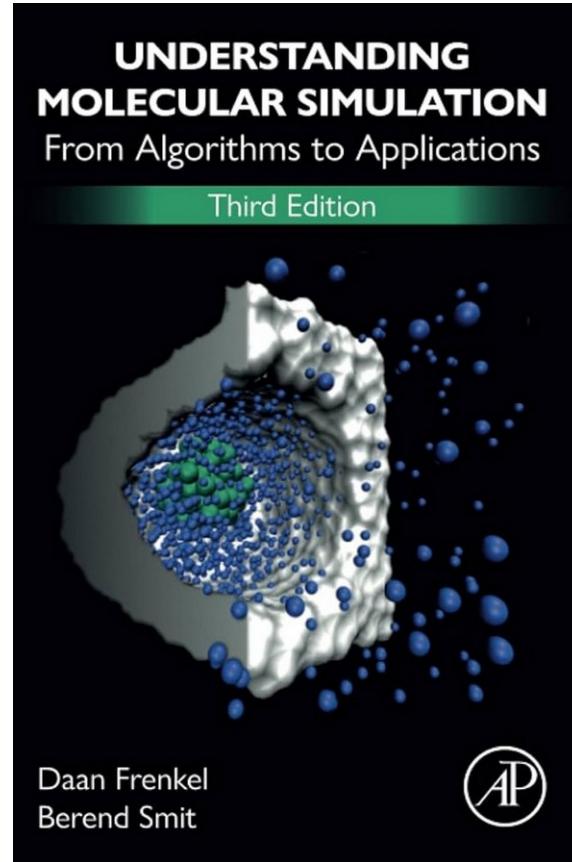
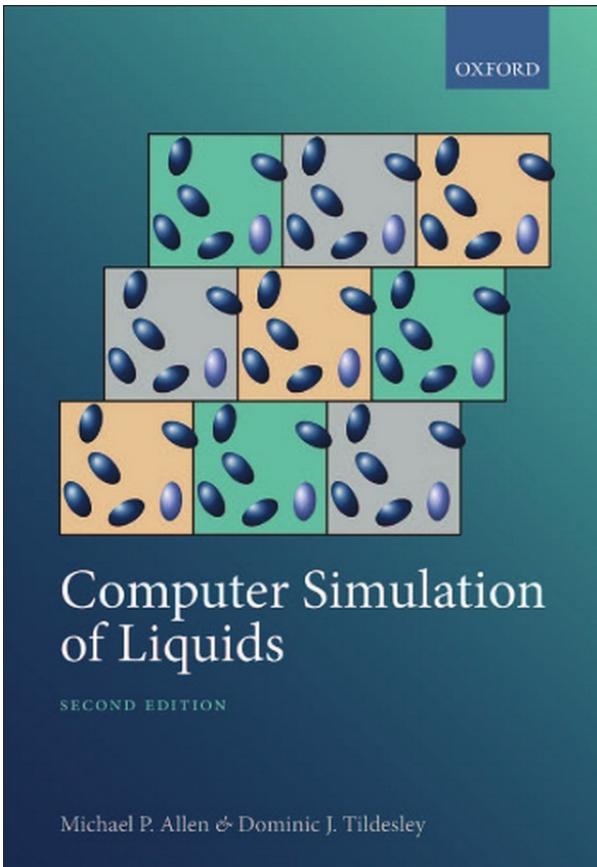
velocity rescaling

Berendsen, Andersen, Nosé-Hoover

Examples of materials science with MD

Accelerated (meta)dynamics

Further resources



Preliminaries: Temperature

So far, we have been working at $T = 0\text{ K}$ (this is true for classical and *ab initio* potentials)

When is temperature important?

When can zero-temperature calculations be used?

Can we get finite T properties using zero-temperature calculations?

Preliminaries: Goals of MD

Thermodynamics: Temperature-dependent materials properties

Ensemble averages of observable quantities

Thermal expansion

Diffusion

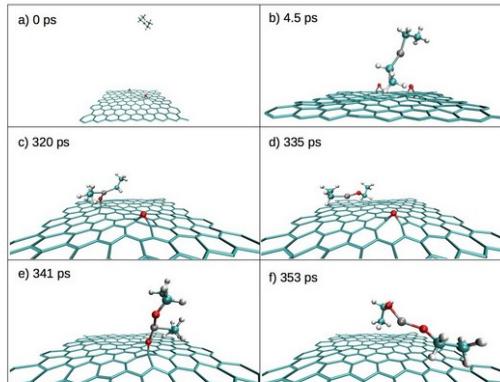
Kinetics

Real-time Dynamics

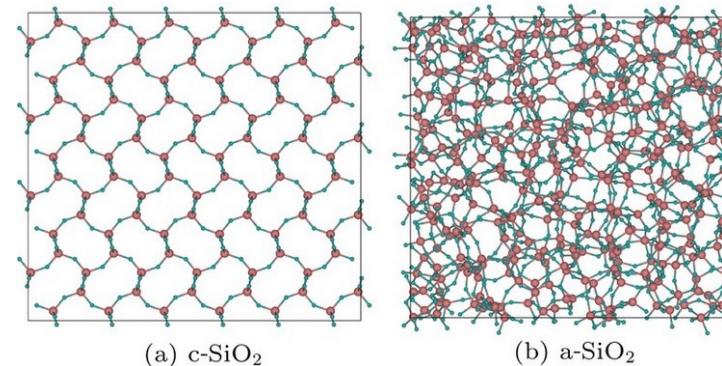
Chemical reactivity (e.g., bond breaking and re-forming)

Catalysis

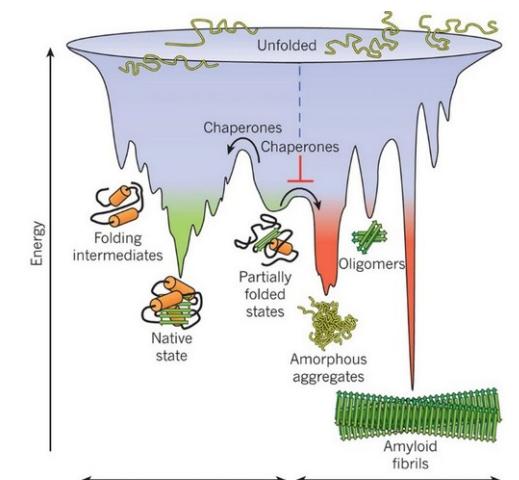
Simulated annealing of complex structures



Schatz Group, Northwestern



Chinese Phys. B 29 027101 (2020)



Nature 475, 324–332 (2011).

Preliminaries: Goals of MD

To capture different phases of matter

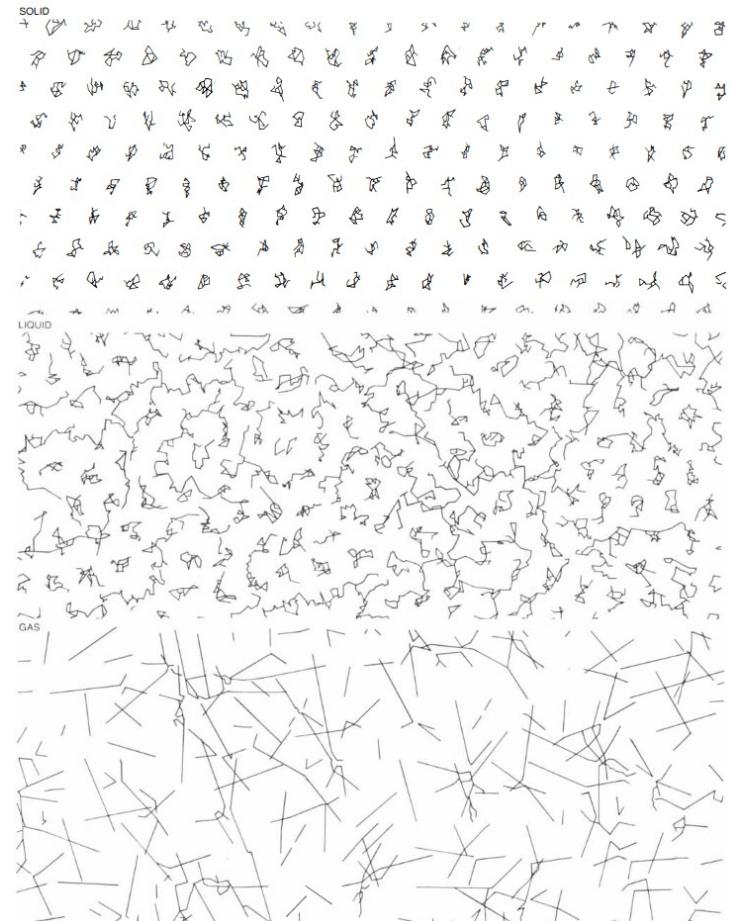
“What is ‘liquid’?” Understanding states of matter

Crystalline Solid: small amplitude vibrations
rare diffusive movements

Liquids: Brownian motion (many collisions)
Diffusive movements

Gas: very long free paths
few collisions

J.A. Barker & D. Henderson. Scientific American, 245, 5, (1981)



TRAJECTORIES OF MOLECULES in a solid, a liquid and a gas were simulated with the aid of a computer. The two-dimensional system of molecules has the same phases and phase transitions as a real substance. The molecular picture is necessarily greatly simplified. In the solid the molecules are constrained to vibrate about fixed lattice sites, whereas molecules in the liquid and the gaseous

phases are free to wander. The only substantial differences between the two fluid states are those of density and of frequency of collision. The computer program calculates the trajectories by solving the equations of motion for some 500 two-dimensional molecules. The simulation was done by Farid Abraham of the International Business Machines Corporation Research Laboratory in San Jose, Calif.

Preliminaries: Temperature

So far, we have been working at $T = 0 \text{ K}$ (this is true for classical and *ab initio* potentials)

When is temperature important?

When can zero-temperature calculations be used?

Can we get finite T properties using zero-temperature calculations?

Temperature: a statistical average of the translational, vibration, rotational motions
of particles (atoms, molecules)

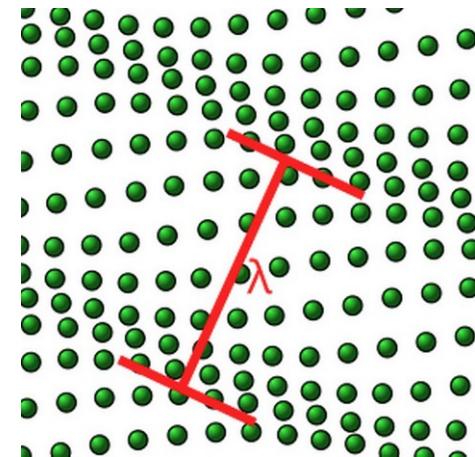
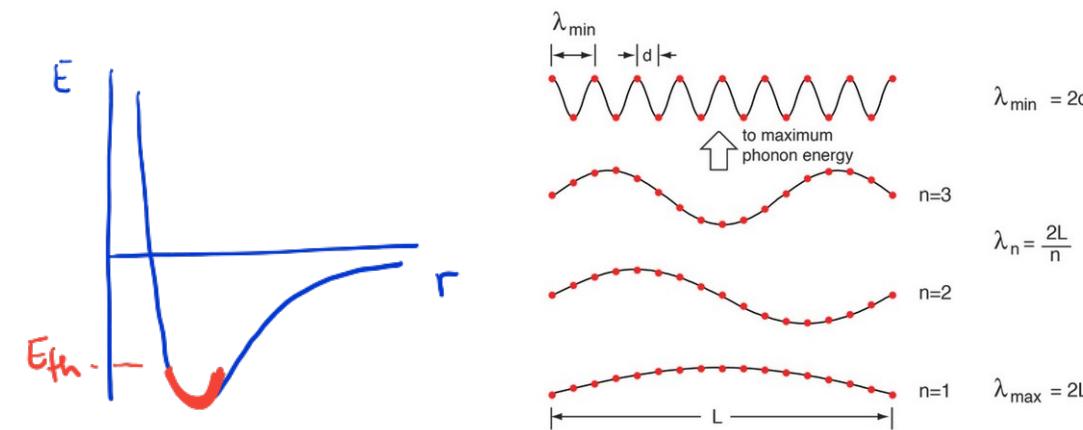
gives way to thermal excitations

$$\left(\frac{\partial E}{\partial T} \right)_V = C_V > 0$$

Examples of T-dependent properties: diffusion, crystal phase changes,
chemical reactivity, electrical and thermal conductivity,
bulk modulus, volume

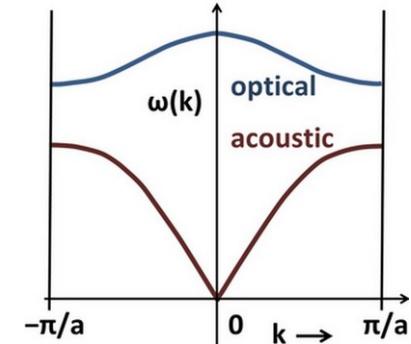
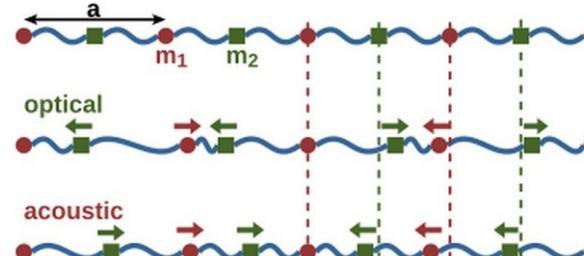
Microscopic picture of thermal excitations

Ionic degrees of freedom: phonons



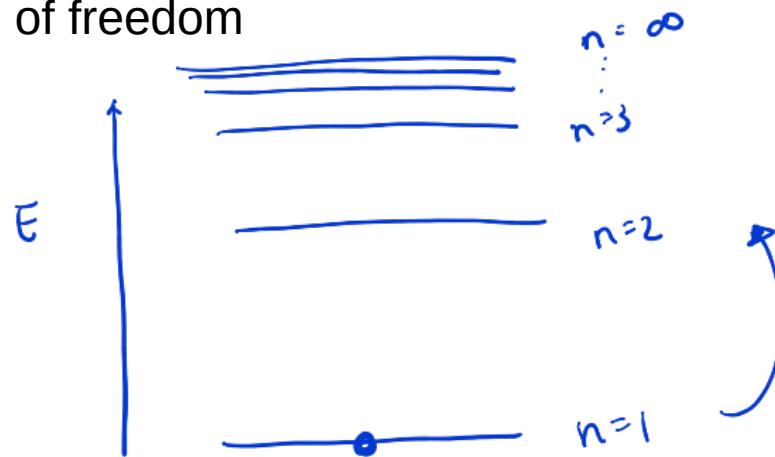
Example of phonons in 1D:

normal modes



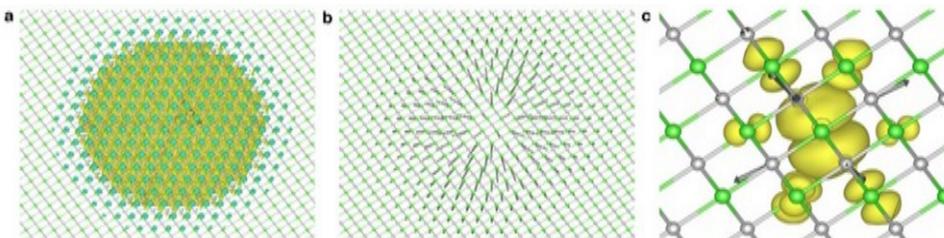
Microscopic picture of thermal excitations

Electronic degrees of freedom



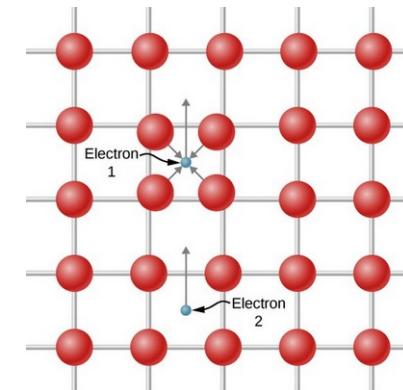
Electron-phonon coupling

Large and small polarons



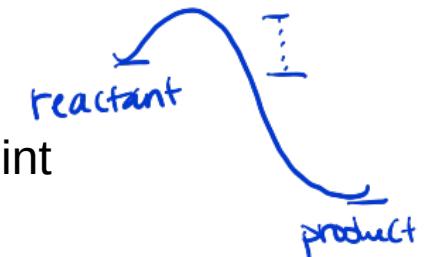
npj Comput Mater 9, 156 (2023).

e.g., in superconductivity
(Cooper pair formation)



Examples where explicit T is not included: Methods

Nudged Elastic Band (NEB) Method: energetic barriers



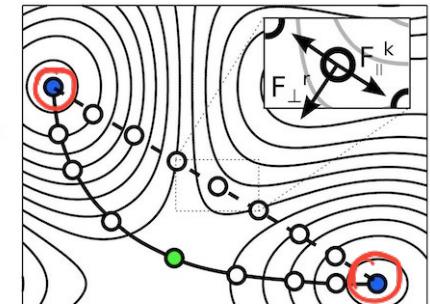
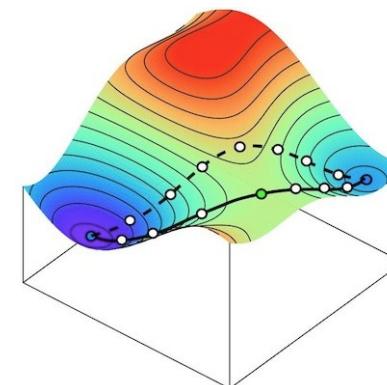
Goal: find the minimum energy path (MEP) through saddle point

- Start with linear path of configurations that are equidistant along reaction coordinate
- Connect configurations or “images” with imaginary springs
- The springs results in a force that nudges each image towards the MEP

Typically use 5-20 images

Need a reasonable guess of
the initial and final states

Energies and forces typically done
with DFT



Examples where explicit T is not included: Methods

Nudged Elastic Band (NEB) Method: energetic barriers

Compute forces **along** and **normal** to the path

$$\mathbf{F}_i^{\text{NEB}} = \mathbf{F}_i^\perp + \mathbf{F}_i^{S\parallel}$$

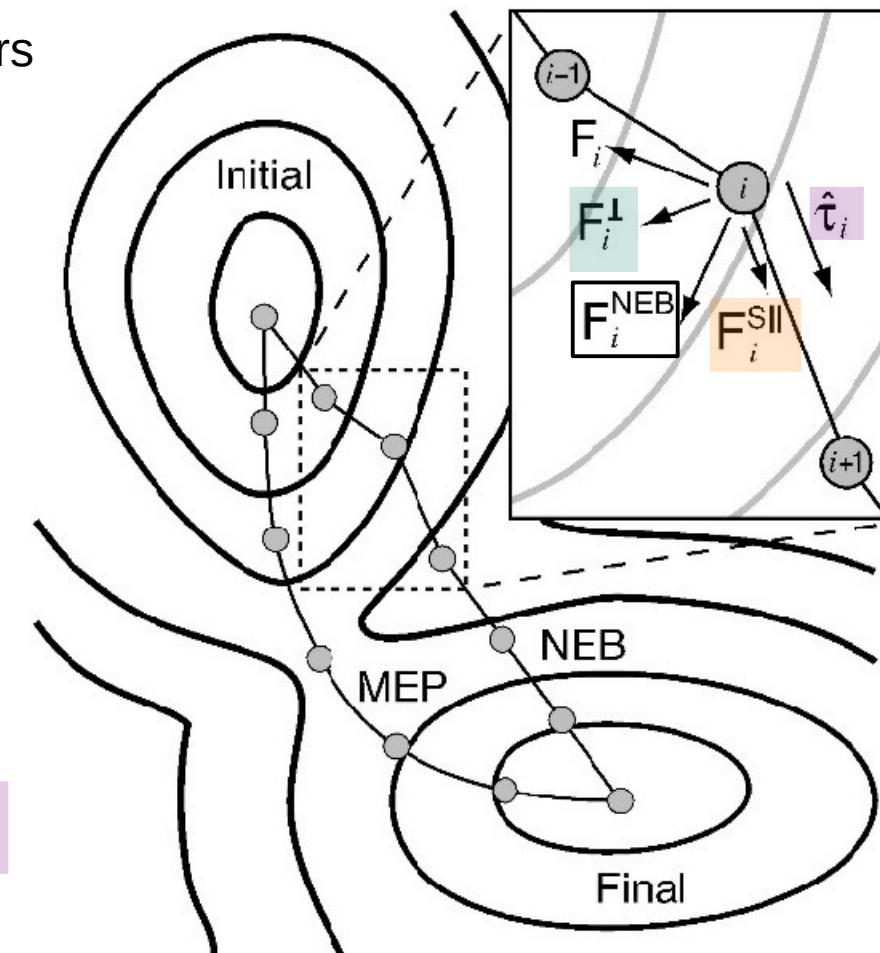
Force from potential

$$\mathbf{F}_i^\perp = -\nabla(\mathbf{R}_i) + \nabla(\mathbf{R}_i) \cdot \hat{\tau}_i$$

Spring force

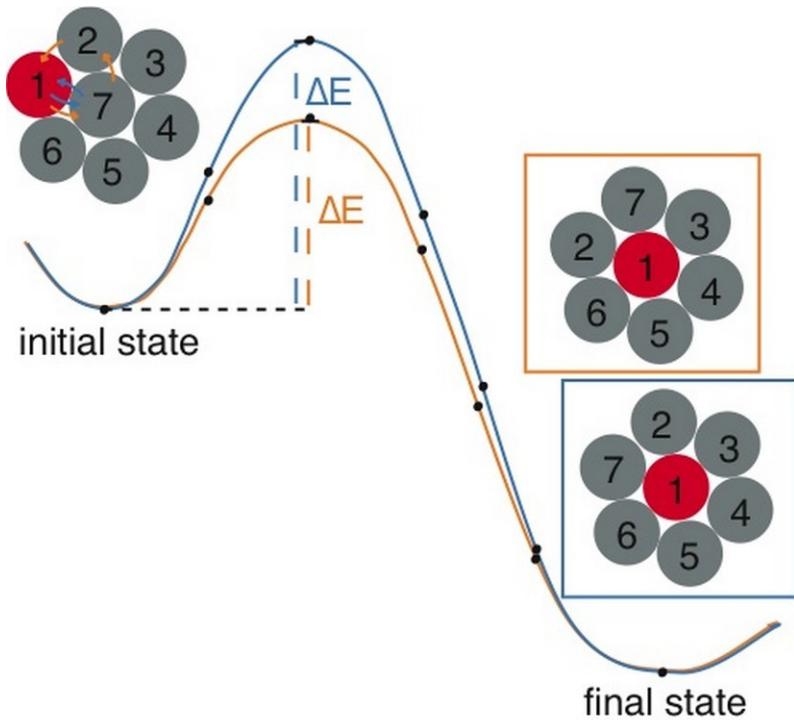
$$\mathbf{F}_i^{S\parallel} = k \left[(\mathbf{R}_{i+1} - \mathbf{R}_i) - (\mathbf{R}_i - \mathbf{R}_{i-1}) \right] \cdot \hat{\tau}_i \hat{\tau}_i$$

tangent $\hat{\tau}_i = \frac{\mathbf{R}_{i+1} - \mathbf{R}_i}{|\mathbf{R}_{i+1} - \mathbf{R}_i|}$



Examples where explicit T is not included: Methods

Energetically-degenerate configurations may be connected via different saddle points

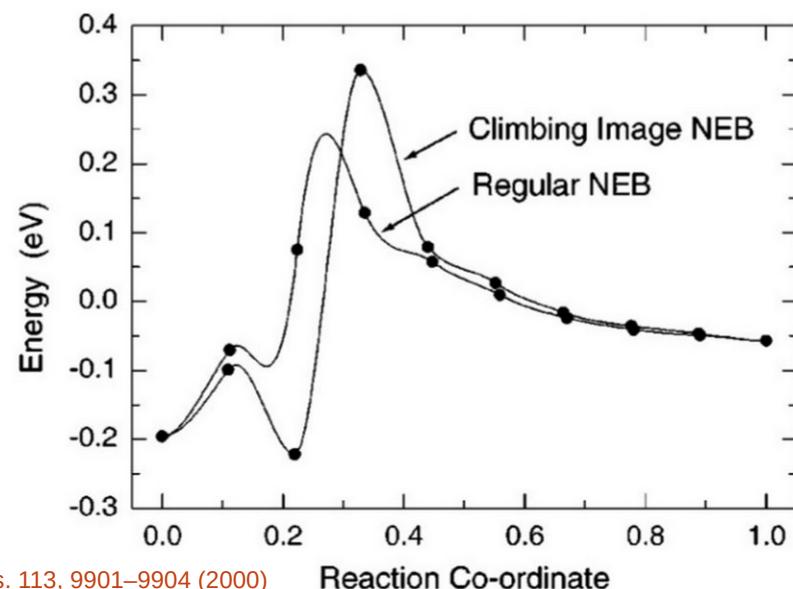


Climbing-image NEB

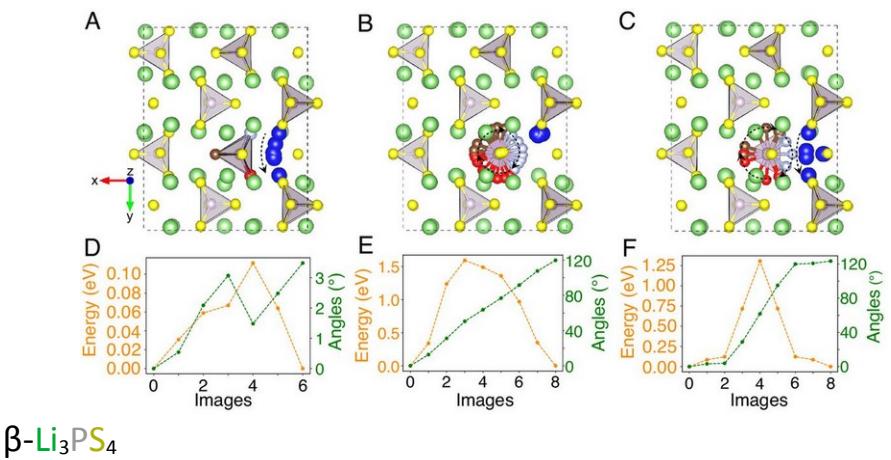
Refine the saddle point

Main change: highest-energy image does not feel spring forces along the band

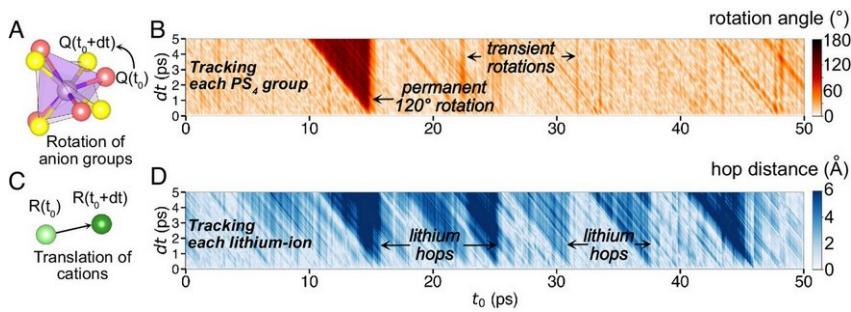
Force is inverted so that this image maximizes its energy along the band and minimizes it in all other directions



Collective motion of lithium with crystal lattice

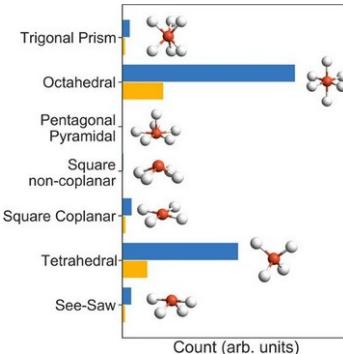
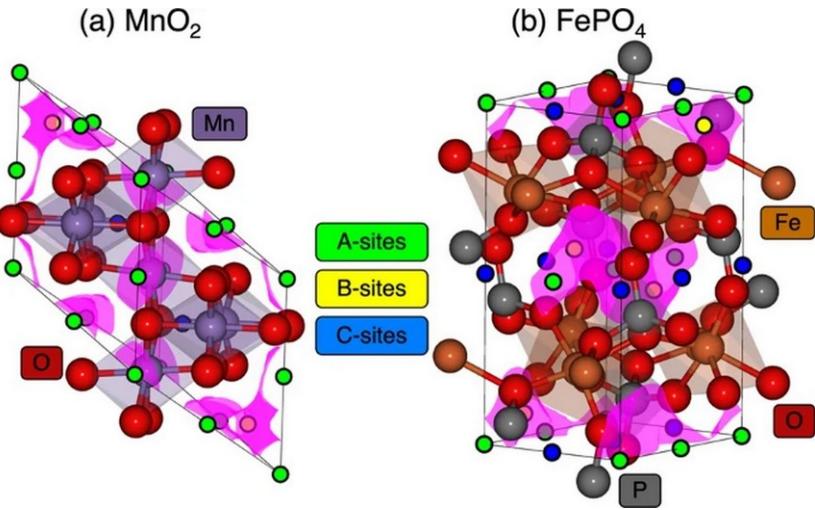


$\beta\text{-Li}_3\text{PS}_4$



PNAS. 121 (18) e2316493121 (2024).

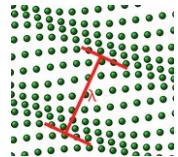
Intercalation of lithium in novel cathode materials



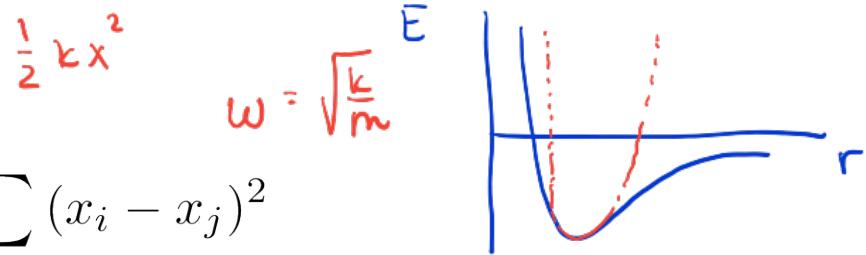
npj Comput Mater 6, 161 (2020).

Examples where explicit T is not included: Methods

Harmonic approximation to lattice vibrations



$$\mathcal{H} = \sum_{i=1}^N \frac{p_i^2}{2m} + \frac{1}{2} m \omega^2 \sum_{ij,nn} (x_i - x_j)^2$$



Main quantity:
interatomic force constants
for along positions τ for direction α/β of atom γ/γ' in cell a/b

Take a Fourier transform

Relation to dynamical matrix

Equations of motion
(generalized eigenvalue problem)

$$C_{\gamma\alpha,\gamma'\beta}(a,b) = \left(\frac{\partial^2 U}{\partial \tau_{\gamma\alpha}^a \tau_{\gamma'\beta}^b} \right)$$

$$\tilde{C}_{\gamma\alpha,\gamma'\beta}(\mathbf{q}) = \sum_b C_{\gamma\alpha,\gamma'\beta}(0,b) e^{i\mathbf{q}\cdot\mathbf{R}_b}$$

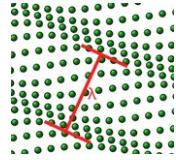
$$\tilde{D}_{\gamma\alpha,\gamma'\beta}(\mathbf{q}) = (M_\gamma M_{\gamma'})^{1/2} \tilde{C}_{\gamma\alpha,\gamma'\beta}(\mathbf{q})$$

$$\sum_{\gamma',\beta} \tilde{C}_{\gamma\alpha,\gamma'\beta}(\mathbf{q}) \underline{u_{m\mathbf{q}}(\gamma'\beta)} = \underline{M_\gamma} \underline{\omega_{m\mathbf{q}}^2} \underline{u_{m\mathbf{q}}(\gamma\alpha)}$$

Phonon displacement pattern Square of phonon frequencies
 Atomic masses

Examples where explicit T is not included: Methods

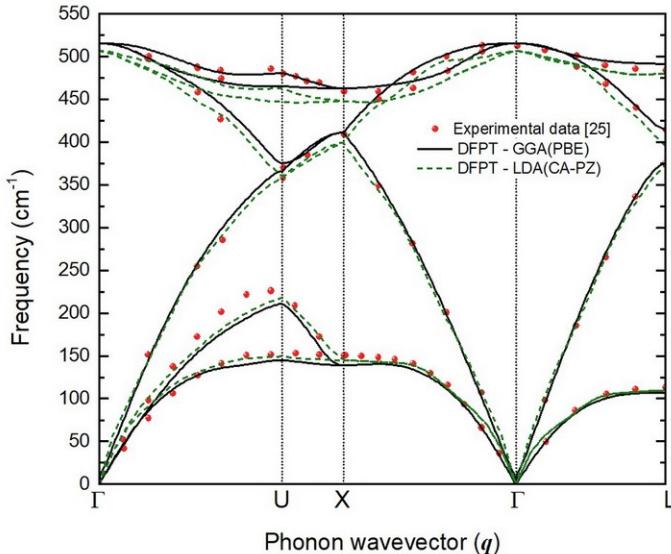
Harmonic approximation to lattice vibrations



$$\mathcal{H} = \sum_{i=1}^N \frac{p_i^2}{2m} + \frac{1}{2} m \omega^2 \sum_{ij,nn} (x_i - x_j)^2$$

e.g., Energies and their second derivatives from DFT/DFPT

Based on total energy differences, can get fairly accurate vibration spectra



Phonon dispersion of crystalline silicon

I González et al 2019 J. Phys.: Condens. Matter 31 425303

Examples where explicit T is not included

As temperature increases, the bulk modulus B decreases
What are the thermodynamic contributions to this change?

$T = 0$

$$B|_{T=0} = V \left(\frac{\partial^2 E}{\partial V^2} \right) \Bigg|_{a_0}$$

Finite T

$$B = V \left(\frac{\partial^2 F}{\partial V^2} \right) \Bigg|_a$$

$$F(V,T) = E_{\text{LDA}}(V) + k_B T \sum_{\mathbf{q},j} \ln \left(2 \sinh \frac{\hbar \omega_{\mathbf{q}j}(V)}{2k_B T} \right)$$

Change in bulk modulus is largely due to change in lattice constant
i.e., small entropic contributions

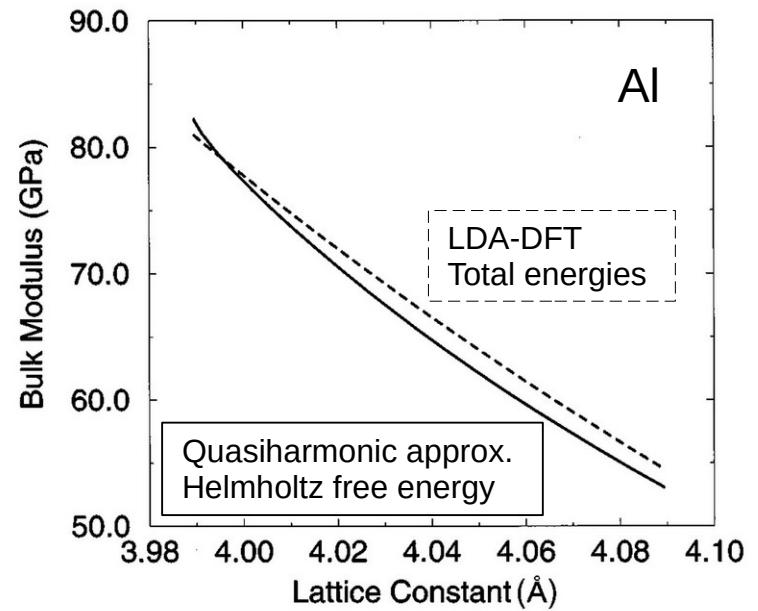


FIG. 2. Calculated bulk modulus as a function of lattice parameter. The solid curve is obtained from the temperature-dependent Helmholtz free energy and the dashed curve is obtained from the $T=0$ electronic energy.

Examples where explicit T is not included

Variation of the coefficient of linear expansion with temperature

$$\alpha(T) = \frac{1}{a_0(T_c)} \left(\frac{da_0(T)}{dT} \right)$$

Two flavors of DFT (LDA, GGA)

Free energy and lattice constant a taken within quasiharmonic approximation again

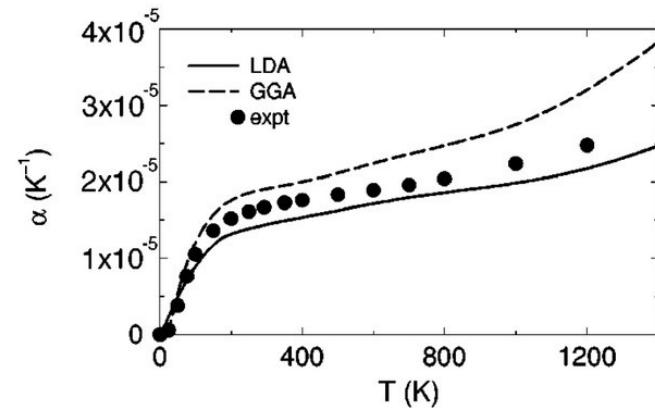


FIG. 2. Coefficient of linear thermal expansion α as a function of temperature. Both the LDA and GGA values are reasonably close to the experimental values; however, the LDA underestimates and the GGA overestimates the thermal expansion. Experimental values are from Ref. 19.

Basics: Governing equations of motion

Basics: Newton's equations

Nuclei are treated classically:

2nd law

$$\bar{F}_i = m_i \bar{a}_i = m_i \frac{d^2 \bar{r}_i}{dt^2} = -\frac{\partial}{\partial \bar{r}_i} U(\bar{r}_1, \dots, \bar{r}_N)$$

$$= -\nabla_i U(\bar{r}_1, \bar{r}_2, \dots, \bar{r}_N) = -\nabla_i U(\bar{r}^N)$$

$$\nabla_i = \left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial y_i}, \frac{\partial}{\partial z_i} \right)$$

For central-force potential

$$U(\bar{r}^N) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \phi_{ij}(r_{ij})$$

$$\bar{F}_i = \sum_{j \neq i} f_{ij} = \sum_{j \neq i} \frac{d}{dr_{ij}} \phi_{ij}(r_{ij}) \underbrace{\frac{\bar{r}_{ij}}{|\bar{r}_{ij}|}}$$

3rd law

$$\bar{f}_{ij} = -\bar{f}_{ji}$$

equations of motion

$$\frac{d^2 \bar{r}_i}{dt^2} = \frac{1}{m_i} \bar{F}_i = \frac{1}{m_i} \sum_{j \neq i} f_{ij} \quad \left. \begin{array}{l} \text{3N coupled differential eq.} \\ \{\bar{r}_i(t), \bar{p}_i(t)\} \rightarrow \{\bar{r}_i(t+\delta t), \\ \bar{p}_i(t+\delta t)\} \end{array} \right\}$$

Basics: Newton's equations

Basics: Verlet algorithm

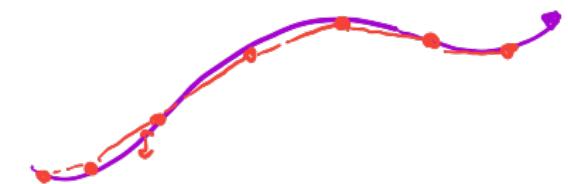
Timestep: δt

Assume \mathbf{a} is constant between t and $t + \delta t$

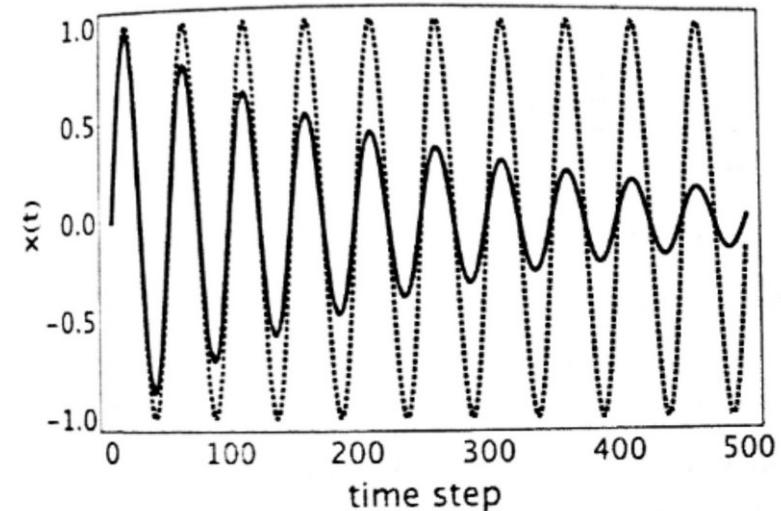
$$\frac{d^2\bar{\mathbf{r}}_i}{dt^2} = \frac{1}{m_i} \sum_{j \neq i} f_{ij}(\mathbf{r}_{ij})$$

$$\bar{a}_i = \frac{d\bar{v}_i}{dt} \quad \bar{v}_i(t + \delta t) = v_i(t) + \int_t^{t + \delta t} \bar{a}_i(t') dt'$$

$$\bar{v}_i = \frac{d\bar{r}_i}{dt} \quad \bar{r}_i(t + \delta t) = \bar{r}_i(t) + \bar{v}_i(t)\delta t + \frac{1}{2}\bar{a}_i(t)\delta t^2$$



Example: 1D Harmonic oscillator



— Naive integration
- - - Exact

Basics: Verlet algorithm

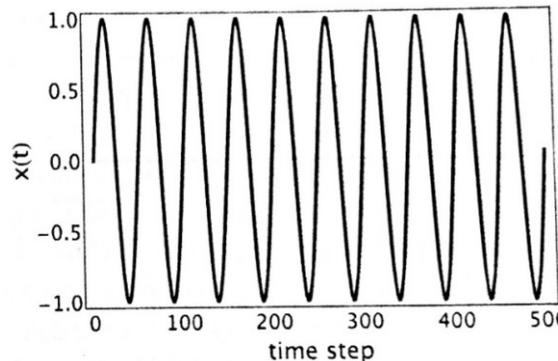
$$\textcircled{1} \quad \mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\delta t + \frac{1}{2}\mathbf{a}_i(t)\delta t^2 + \frac{1}{6}\bar{\mathbf{b}}_i(t)\delta t^3 + \mathcal{O}(\delta t^4)$$

$$\textcircled{2} \quad \bar{\mathbf{r}}_i(t - \delta t) = \mathbf{r}_i(t) - \mathbf{v}_i(t)\delta t + \frac{1}{2}\mathbf{a}_i(t)\delta t^2 - \frac{1}{6}\bar{\mathbf{b}}_i(t)\delta t^3 + \mathcal{O}(\delta t^4)$$

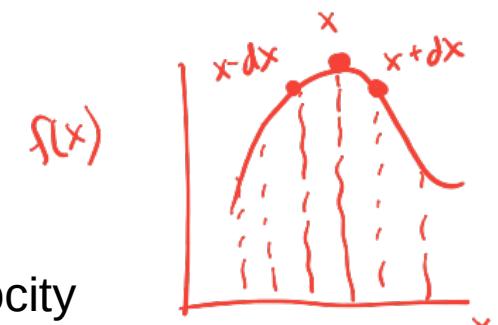
$$\textcircled{1} + \textcircled{2} \quad \bar{\mathbf{r}}_i(t + \delta t) = 2\bar{\mathbf{r}}_i(t) - \bar{\mathbf{r}}_i(t - \delta t) + \bar{\mathbf{a}}_i(t)\delta t^2 \quad \left. \begin{array}{l} \text{not relevant to} \\ \text{forces/energies} \end{array} \right\}$$

$$\textcircled{1} - \textcircled{2} \quad \mathbf{v}_i(t) = \frac{\bar{\mathbf{r}}_i(t + \delta t) - \bar{\mathbf{r}}_i(t - \delta t)}{2\delta t} \quad \left. \begin{array}{l} K_i := \frac{1}{2}m_i v_i^2 \end{array} \right\}$$

"non-self starting"



Verlet: 2nd order in velocity
4th order in position



Basics: Verlet algorithm

Velocity Verlet Algorithm "self-starting" → same value of time

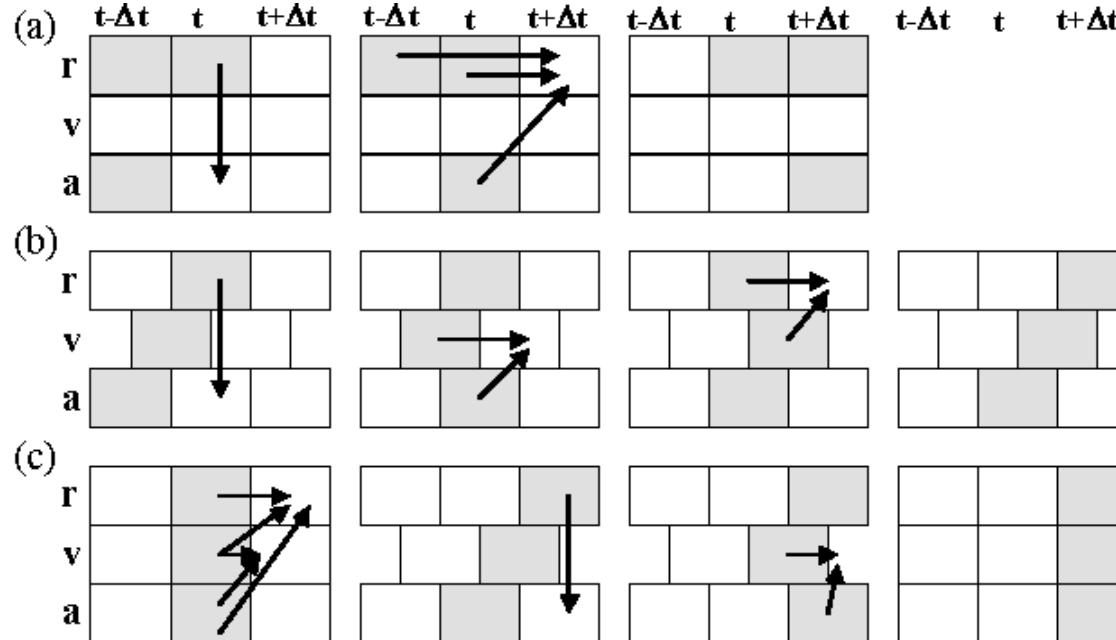
$$\bar{r}_i(t + \delta t) = \bar{r}_i(t) + v_i(t) \delta t + \frac{1}{2} \bar{a}_i(t) \delta t^2$$

$$\bar{v}_i(t + \delta t) = v_i(t) + \frac{1}{2} [a_i(t) + a_i(t + \delta t)] \delta t$$

$$\bar{a}_i(t) = \frac{\bar{F}_i(t)}{m} = \frac{1}{m} \nabla U(\bar{r}_i(t))$$

1. Calculate $\mathbf{v}(t + \frac{1}{2} \Delta t) = \mathbf{v}(t) + \frac{1}{2} \mathbf{a}(t) \Delta t.$
2. Calculate $\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \mathbf{v}(t + \frac{1}{2} \Delta t) \Delta t.$
3. Derive $\mathbf{a}(t + \Delta t)$ from the interaction potential using $\mathbf{x}(t + \Delta t).$
4. Calculate $\mathbf{v}(t + \Delta t) = \mathbf{v}(t + \frac{1}{2} \Delta t) + \frac{1}{2} \mathbf{a}(t + \Delta t) \Delta t.$

Basics: Verlet algorithm



- (a) original Verlet algorithm
- (b) leap-frog Verlet algorithm
- (c) velocity form of Verlet algorithm

Grey boxes indicate stored variables

Basics: Other ODE integration schemes

Predictor-corrector, e.g., Gear

Predictor: from known positions and their time derivatives up to order q

Predict same quantities at $t+\Delta t$ via Taylor expansion
including the “predicted acceleration”, \mathbf{a}

Force evaluation: force from gradient of potential at *predicted* positions
acceleration from these forces, “gradient acceleration” \mathbf{a}^*
compute “error signal” = $\mathbf{a}^* - \mathbf{a}$

Corrector: use the “error signal” to “correct” positions and their time derivatives
corrections are proportional to the “error signal” via a coefficient determined
to maximize the stability of the algorithm

Read more [here](#)

Basics: Other ODE integration schemes

Predictor-corrector, e.g., Gear

1. Predict positions and velocities at the next timestep

$$\begin{aligned} r(t + \delta t) &= r(t) + \dots & a(t + \delta t) &= a(t) + \dots \\ v(t + \delta t) &= v(t) + \dots & b(t + \delta t) &= b(t) + \dots \end{aligned}$$

2. Evaluate forces at the next step using predicted positions

Error between the predicted and newly calculated acceleration

$$\Delta a(t + \delta t) = a^c(t + \delta t) - a^p(t + \delta t)$$

3. Correct the predicted positions and velocities

$$\begin{aligned} r^c(t + \delta t) &= r^p(t + \delta t) + c_0 \Delta a(t + \delta t) & a^c(t + \delta t) &= a^p(t + \delta t) + c_2 \Delta a(t + \delta t) \\ v^c(t + \delta t) &= v^p(t + \delta t) + c_1 \Delta a(t + \delta t) & b^c(t + \delta t) &= b^p(t + \delta t) + c_3 \Delta a(t + \delta t) \end{aligned}$$

Coefficients c_i are chosen to maximize stability and depend on the algorithm chosen

Read more [here](#)

Basics: Conservation laws

Hamiltonians : Newton's eqns in terms of energies

$$H(\vec{p}^N, \vec{r}^N) = K(\vec{p}^N) + U(\vec{r}^N)$$

" " $\vec{r}^N = \vec{r}_1, \vec{r}_2, \vec{r}_3, \dots, \vec{r}_N$

* conservation of energy

$$\frac{dH}{dt} = 0$$

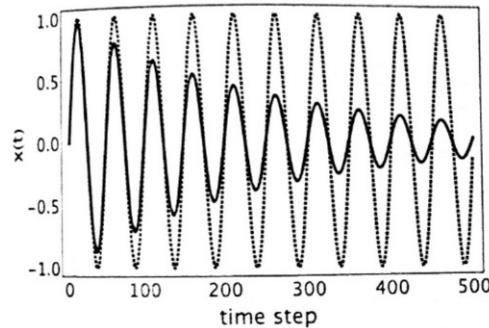
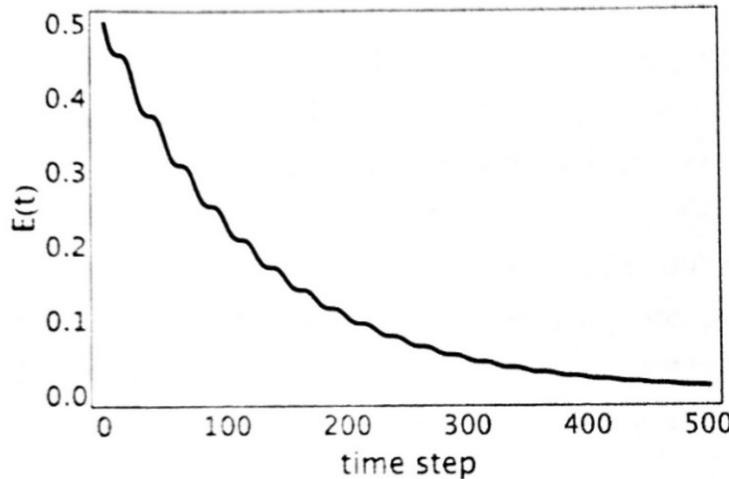
i.e., potential energy
not depend

velocities
(i.e., friction)

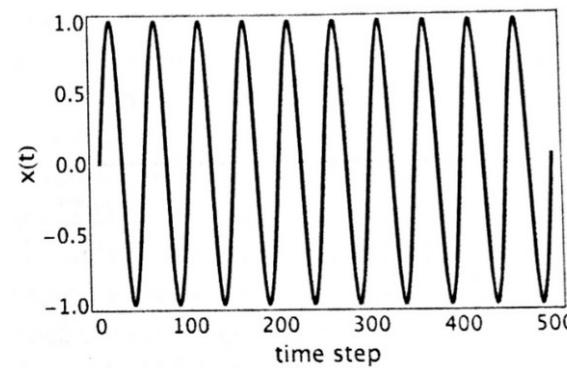
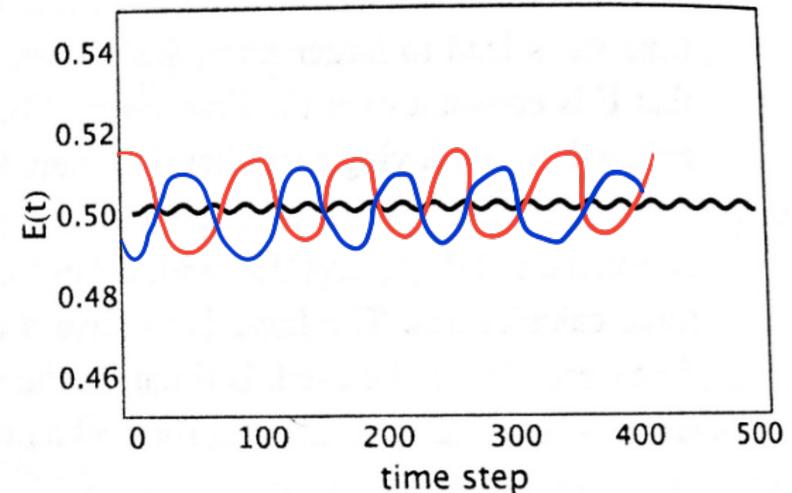
$$\begin{aligned}\frac{d}{dt} H &= \frac{d}{dt} \left[\frac{1}{2}mv^2 + U(x) \right] \\ &= mv \frac{dv}{dt} + \frac{du}{dt} = \cancel{mva} + \underbrace{\frac{du}{dx}}_{-F} \underbrace{\frac{dx}{dt}}_v = mva - mav = 0\end{aligned}$$

Basics: Conservation laws, example

non-Verlet



Verlet



Basics: Conservation laws, example

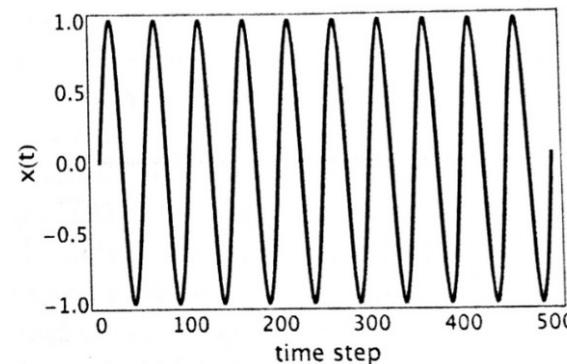
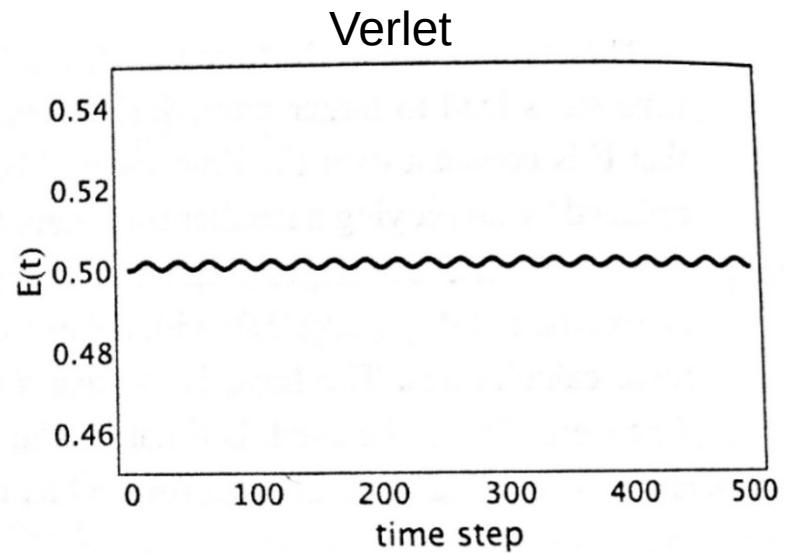
A note on (numerical) fluctuations

due to inexact integration scheme
of eqns of motion

rule of thumb $\frac{\max E - \min E}{\langle E \rangle} \sim 10^{-4}$

poor stability of calc

- drift over time
- explosion of energies
- large oscillation



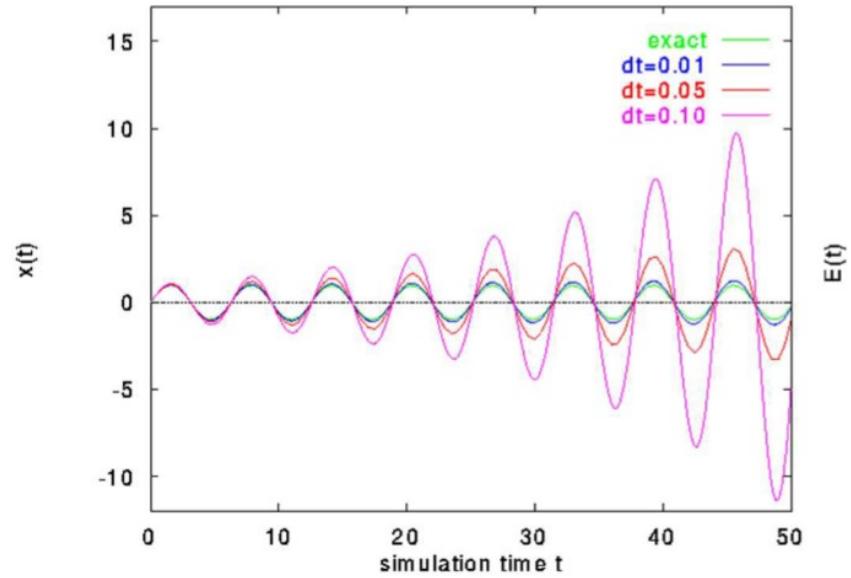
Basics: Conservation laws, example

choosing δt

Typical for condensed matter systems: 0.5 – 10 fs
determined by relevant time scale to system,
e.g., shortest vibrational mode
e.g., for liquid, mean time b/t collisions ~ 10 fs
Typical trajectory length: $>\sim 10$ ps

Good rules of thumb for orders of magnitude

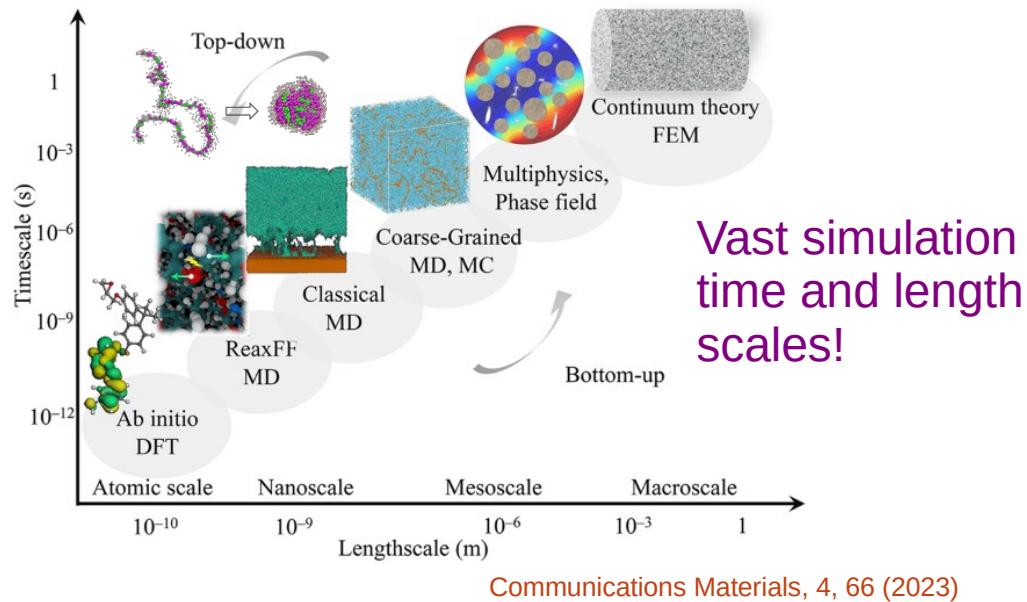
Actual value may depend on the problem
Do relevant convergence tests (on a small, representative system)!



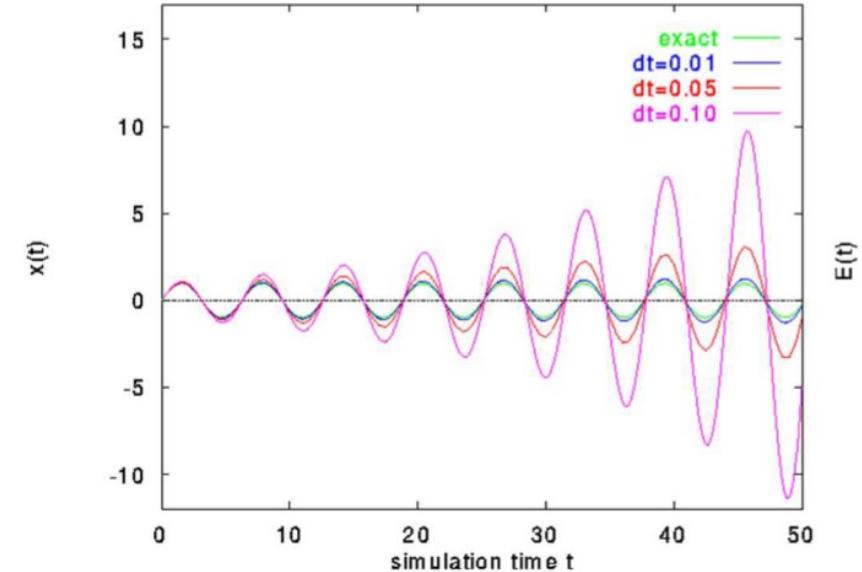
B. Kozinsky, APCOMP 275

Basics: Conservation laws, example

choosing δt



Typical for condensed matter systems: 0.5 – 10 fs
determined by relevant time scale to system,
e.g., shortest vibrational mode
e.g., for liquid, mean time b/t collisions ~ 10 fs
Typical trajectory length: $>\sim 10$ ps



$$10^{-15} \text{ s}$$

Basics: where do the forces come from?

Inc. accuracy

Forces obtained from a potential defined through:

- Pair potentials (e.g., Lennard-Jones, Morse)
- Multi-body potentials (e.g., EAM)
- Reactive potentials (e.g., ReaxFF)
- Tight-binding methods (semi-empirical QM)
- Quantum mechanics (e.g., DFT, HF)

Inc. complexity
and cost

Calculating forces is expensive:

Derivatives are numerically tricky

Depends on nature of the potential:

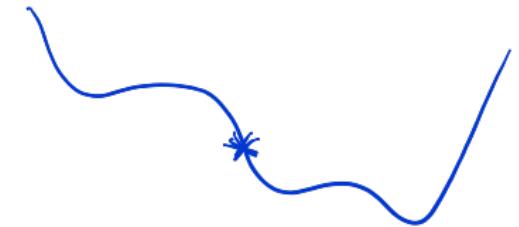
Analytical derivatives are best

Short- v long-range interactions

Basics: Overview of MD calculation

Initialize: choose simulation cell size/shape, boundary conditions
choose positions and momenta

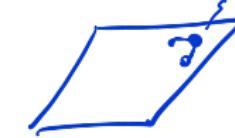
Integrate: propagate equations of motions
compute forces, determine new positions



Equilibrate: allow system to reach a stable state (i.e., lost memory of initial conditions)
Keep track of energies and other relevant thermodynamic quantities

Collect statistics: accumulate quantities of interest; averages, correlations, etc.

Basics: Initialization



Choose simulation cell shape and size:

symmetric cells are often easier to work with, less chance of strange image effects
minimize finite-size effects! e.g., supercells



Initialize positions: determined by problem you want to simulate;
start from something you know is stable or close to problem of interest

e.g., from a (perturbed) crystal structure
e.g., somewhat random positions (e.g., liquids)

avoid short distances (high forces and energies!) and high velocities
sometimes a bit of trial-and-error is involved



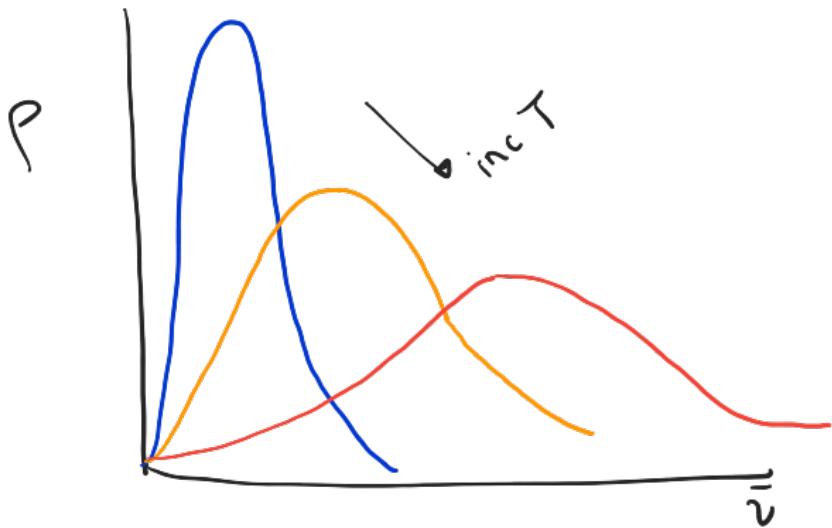
Be careful of codes that constrains the calculation to symmetry!

Basics: Initialization

Initialize velocities:

From Maxwell-Boltzmann distribution

$$\rho(\mathbf{v}_i) \approx \left(\frac{m}{2\pi k_B T} \right)^{\frac{3}{2}} v_i^2 \exp \left(- \frac{mv_i^2}{2k_B T} \right)$$



$$\rho(v_{ix}, v_{iy}, v_{iz}) = \left(\frac{m_i}{2\pi k_B T} \right)^{\frac{3}{2}} \exp \left(- \frac{(p_{ix}^2 + p_{iy}^2 + p_{iz}^2)}{2mk_B T} \right)$$

$$v_{\text{rms}} = \sqrt{\frac{3k_B T}{m}}$$

Generating velocities randomly from Gaussian distribution
Box & Muller (1958)

$$\sigma = \sqrt{\frac{k_B T}{m}} \quad \rho(v) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left(- \frac{(v - \langle v \rangle)^2}{2\sigma^2} \right)$$

random $v \rightarrow$

$v = \langle v \rangle + \sigma \tau$

choose $\langle v \rangle = 0$

$\tau_1 = (-2 \ln \mathcal{R}_1)^{\frac{1}{2}} \cos(2\pi \mathcal{R}_2)$

$\tau_2 = (-2 \ln \mathcal{R}_1)^{\frac{1}{2}} \sin(2\pi \mathcal{R}_2)$

$\mathcal{R}_i \in (0, 1)$

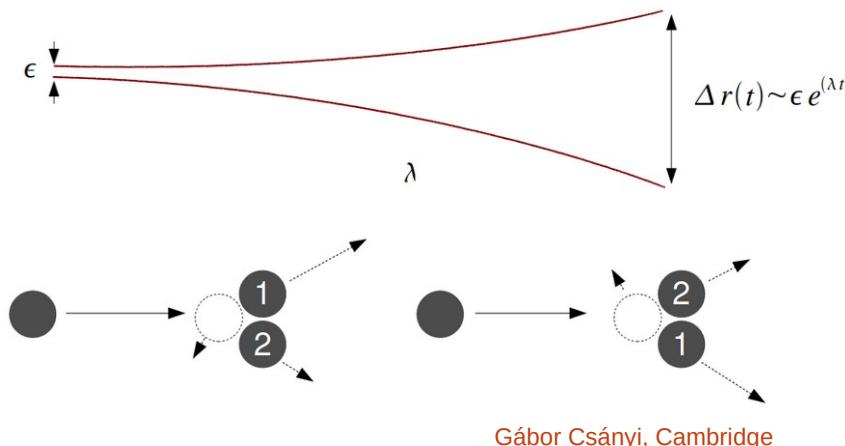
Basics: Initialization

Lyapunov instability

Trajectories sensitive to initial conditions

Butterfly effect:

Two trajectories will diverge exponentially as time progresses



Not an issue if collect **sufficient statistics** rather than predicting the future

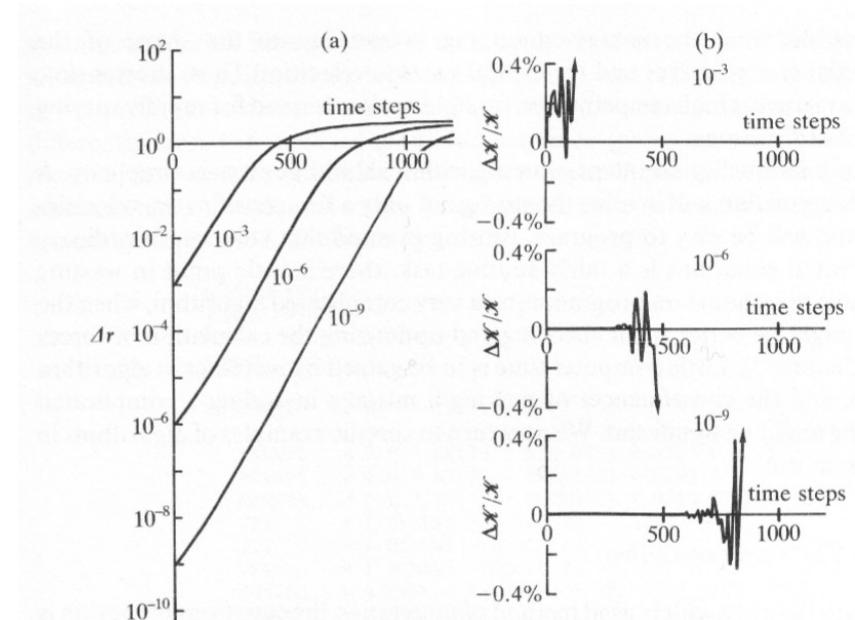


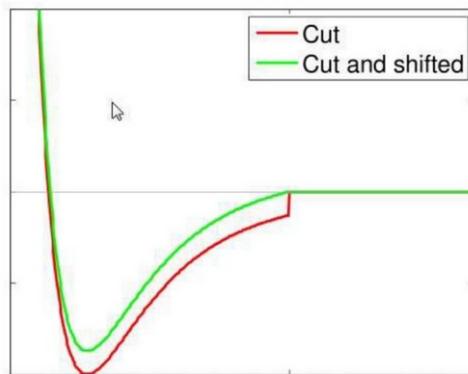
Fig. 3.1 The divergence of trajectories in molecular dynamics. Atoms interacting through the potential $v^{\text{RLJ}}(r)$, eqn (1.10a), were used, and a dense fluid state was simulated ($\rho^* = 0.6$, $T^* = 1.05$, $\delta t^* = 0.005$). The curves are labelled with the initial displacement in units 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.

Allen & Tildesley

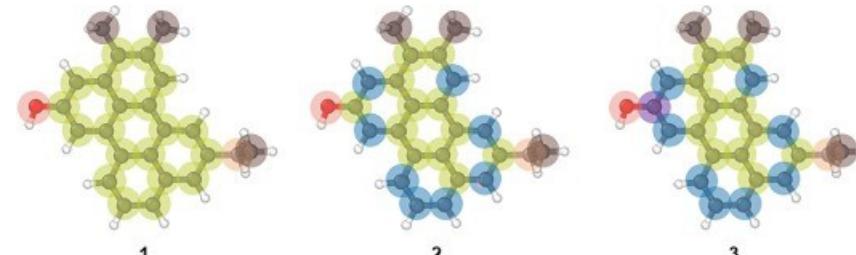
Basics: Minimizing computational cost

At minimum, two-body potentials are involved in evaluations of forces (and positions):

- **Neighbor lists:** generate list of pairs of atoms to compute over
update list when atoms move outside r_c
- **Lookup tables:** interpolate pre-computed interaction energies at various distances
- **Periodic boundary conditions:** simulate an infinite system with finite number of particles
- **Potential cutoffs:**



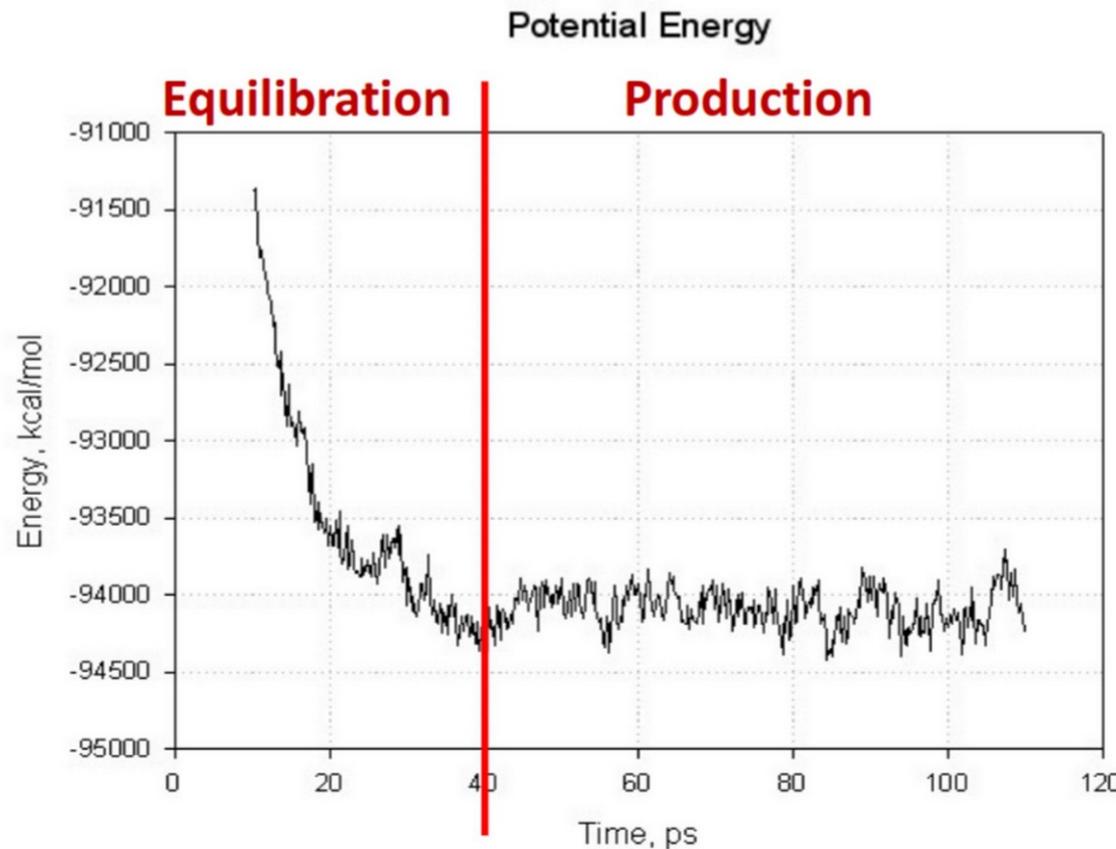
- **Remove degrees of freedom:**
e.g., remove fast vibrations → increase time step
useful for polymers and macromolecules



Assaf et al. Materials & Design. 240, 112831 (2014)

See also: J.G. Gay and B.J. Berne, Journal of Chemical Physics, 74, 3316 (1981)

Equilibration



An example: Lennard-Jonesium

$$\phi(r_{ij}) = \frac{B}{r_{ij}^{12}} - \frac{A}{r_{ij}^6} = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right]$$

$$\epsilon = A^2/4B$$
$$\sigma = (B/A)^{1/6}$$

$\phi(\sigma) = 0$

Reduced units

$$\phi^*(r^*) = 4 \left[\left(\frac{1}{r^*} \right)^{12} - \left(\frac{1}{r^*} \right)^6 \right]$$

$$\phi^* = \phi/\epsilon$$
$$r^* = r/\sigma$$

Table 6.1 Reduced units in Lennard-Jones systems

Value	In reduced units
Potential energy	$U^* = U/\epsilon$
Temperature	$T^* = k_B T/\epsilon$
Density	$\rho^* = \rho \sigma^3$
Pressure	$P^* = P \sigma^3/\epsilon$
Time	$t^* = t/t_o$, where $t_o = \sigma \sqrt{m/\epsilon}$

ϵ and σ are defined in Eq. (6.21). All energies are in units of ϵ , e.g., $E^* = E/\epsilon$, $K^* = K/\epsilon$.

An example: Lennard-Jonesium

Forces (analytical!):

$$\mathbf{f}_{ij}^* = \frac{d\phi^*}{dr^*} = -24 \frac{\mathbf{r}^*}{r^{*2}} \left[\left(\frac{2}{r^{*2}} \right)^{12} - \left(\frac{1}{r^{*2}} \right)^6 \right]$$

Potential cutoffs:

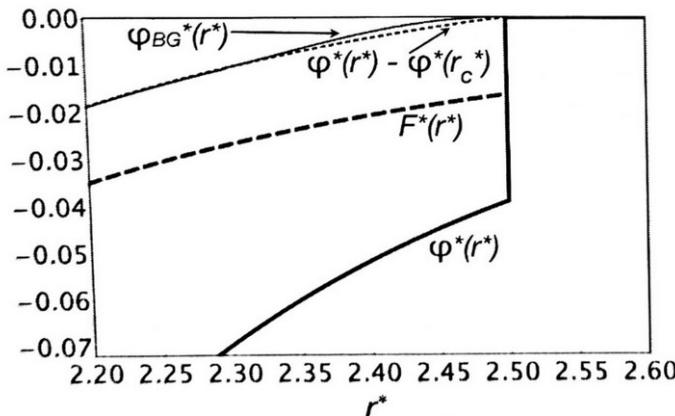


Figure 6.4 Force F^* , potential ϕ^* , and shifted potential $\phi^*(r^*) - \phi^*(r_c^*)$ for an LJ potential cutoff at $r_c = 2.5\sigma$. Note that there is a discontinuity in both the potential and force at the cutoff. $\phi_{BG}^*(r^*)$ is a potential that smoothly interpolates the potential and derivative to zero at $r_c^* = 2.5$, avoiding the discontinuities [43].

An example: Lennard-Jonesium

Equilibration:

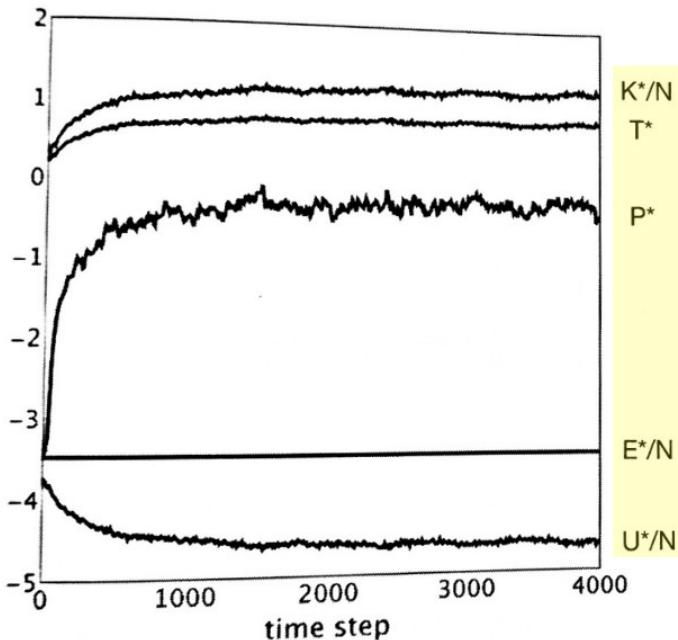


Figure 6.5 Initial equilibration of the Lennard-Jones system described in the text. Shown are instantaneous values for K^*/N , T^* , P^* , E^* , and U^*/N . Note that while K^* and U^* vary greatly over the course of the first 1500 time steps or so, the energy E^* is constant.

An example: Lennard-Jonesium

Equilibration:

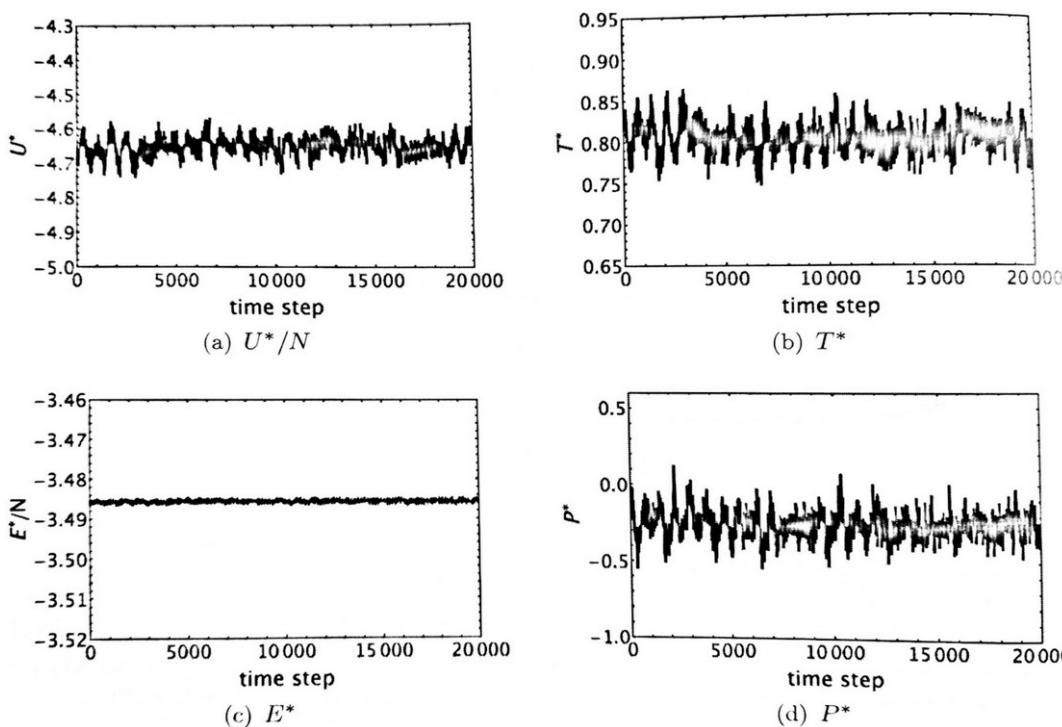


Figure 6.6 Instantaneous values in an MD simulation of the Lennard-Jones system from Figure 6.5.

$$\Delta(U^*/N) \sim 0.17$$

$$\Delta(E^*/N) \sim 0.0009$$

$$\frac{\Delta(E^*/N)}{\langle E^*/N \rangle} \sim 0.0003$$

An example: Lennard-Jonesium

Analysis of trajectories: Averaging of quantities, defined as

$$\langle O \rangle = \frac{1}{\tau_{max}} \sum_{\tau=1}^{\tau_{max}} O(\tau)$$

Avoid correlations by periodic sampling or binning

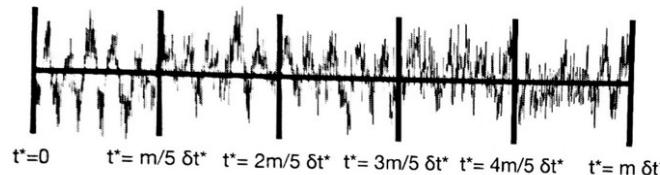


Figure 6.7 Schematic of binning procedure. The total number of time steps (m) is broken into n_b sections and averaged over each of those sections.

Bin width should be selected to lead to uncorrelated results
Balance between enough time steps to be uncorrelated and
enough bins so that bin statistics are meaningful

Average over bins

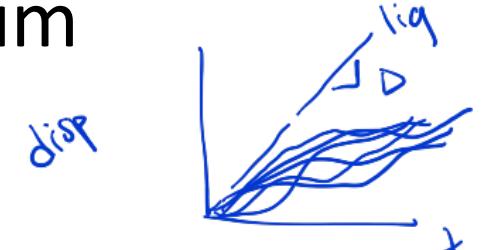
$$\langle U \rangle = \frac{1}{n_b} \sum_{i=1}^{n_b} \langle U \rangle_i$$

Average within bins

$$\langle U \rangle_i = \frac{1}{\tau_b} \sum_{\tau=\tau_o}^{\tau_o+\tau_b} U(\tau)$$

An example: Lennard-Jonesium

Analysis of trajectories: Variance of quantities



$$\sigma_U^2 = \frac{1}{\tau_{max}} \sum_{\tau=1}^{\tau_{max}} [U(\tau) - \langle U \rangle]^2 = \langle [U(\tau) - \langle U \rangle]^2 \rangle = \langle U^2 \rangle - \langle U \rangle^2$$

$$\frac{\sigma_U}{\langle U \rangle} \propto \frac{1}{\sqrt{N}}$$

Relative error in average of U scales inversely with system size

$\sigma(\langle U \rangle_i^2)$ Variance of mean of bin averages of U ~ statistical measure of quality of data
i.e., error bars

See also Allen & Tildesley

S.K. Schiferl & Duane C. Wallace. J. Chem. Phys. 83, 5203-5209 (1985)

NB: $\sigma^2 = \frac{1}{n} \sum_i^n (x_i - \bar{x})^2$ vs $\sigma^2 = \frac{1}{n-1} \sum_i^n (x_i - \bar{x})^2$

Sample variation as an unbiased estimation of the population variance

An example: Lennard-Jonesium

Analysis of trajectories: Correlations of quantities

↪ degree to which the response of one variable
is influenced/coupled to another

$$C_{AB} = \frac{\langle (A - \langle A \rangle)(B - \langle B \rangle) \rangle}{\sigma_A \sigma_B} \quad \sigma_A = \langle (A - \langle A \rangle)^2 \rangle$$

$$\equiv \begin{cases} 1 & \text{perfectly correlated} \\ 0 & \text{independent} \end{cases}$$

An example: Lennard-Jonesium

Analysis of trajectories: Correlations of quantities

Spatial correlation functions: **radial distribution function**

Lennard-Jones potential: 10,976 atoms, NVE ensemble

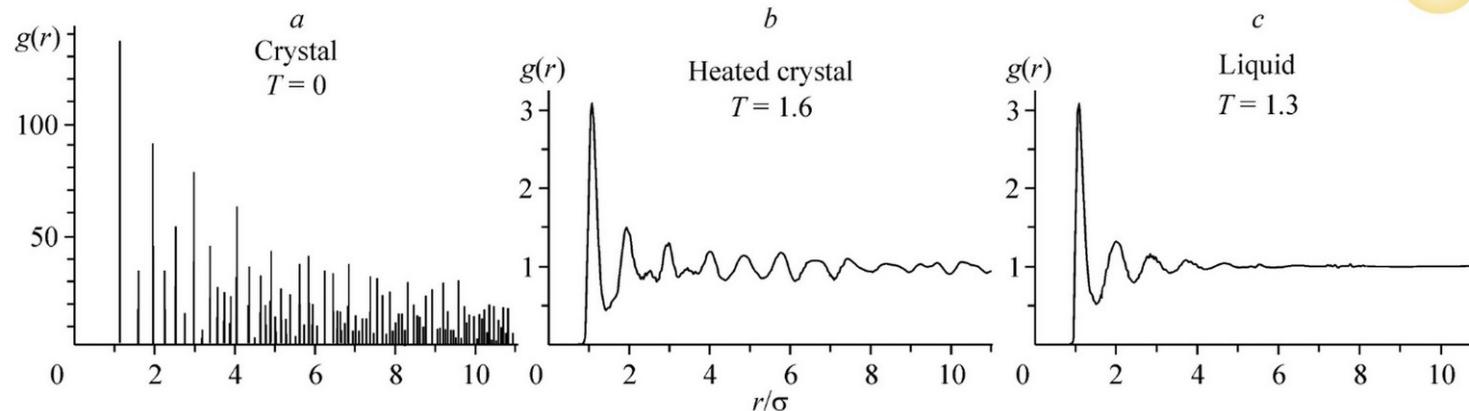
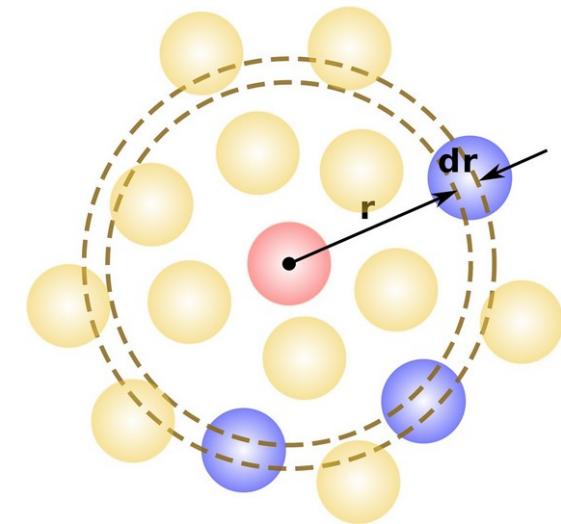


Fig. 3. Radial distribution function of an fcc crystal at $T = 0$ (a), of a crystal heated to $T = 1.6$ (b), after melting (c).



How do
we know
the solid
is melted?

An example: Lennard-Jonesium

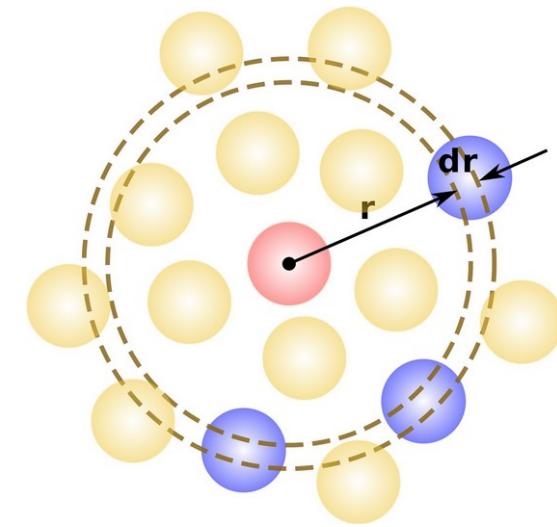
Analysis of trajectories: Correlations of quantities

Spatial correlation functions: **radial distribution function**

"probability of finding a pair of particles $i + j$
a distance r_{ij} apart"

$$g(r) = \frac{V}{4\pi r^2 N^2} \left\langle \sum_i \sum_{j \neq i} \delta(\vec{r} - \vec{r}_{ij}) \right\rangle$$
$$\underbrace{\frac{1}{4\pi r^2} \cdot \frac{1}{N}}$$

number density = $\rho = \frac{N}{V}$



An example: Lennard-Jonesium

Analysis of trajectories: Correlations of quantities

Spatial correlation functions: **radial distribution function**

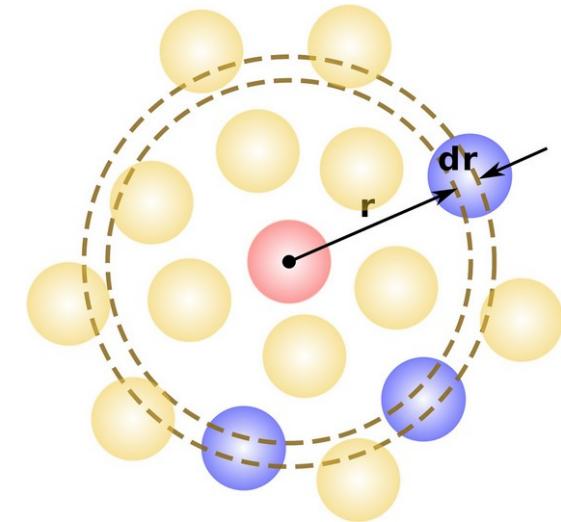
From $g(r)$, compute average number of neighbors

$$n = 4\pi \int_0^{r_0} r^2 g(r) dr$$

Using $g(r)$, rewrite average potential energy & pressure

$$\langle U \rangle = 2\pi \frac{N^2}{V} \int_0^\infty r^2 \phi(r) g(r) dr$$

$$\langle P \rangle = Nk_B T - \frac{2\pi}{3} \frac{N^2}{V} \int_0^\infty r^2 \left(r \frac{\partial \phi(r)}{\partial r} \right) g(r) dr$$



An example: Lennard-Jonesium

Analysis of trajectories: Correlations of quantities

Spatial correlation functions: **translational order parameter**

scaled density $\rho^* = 0.55$

$\tilde{\rho}(\vec{k})$

$\rho^* = 1.0$

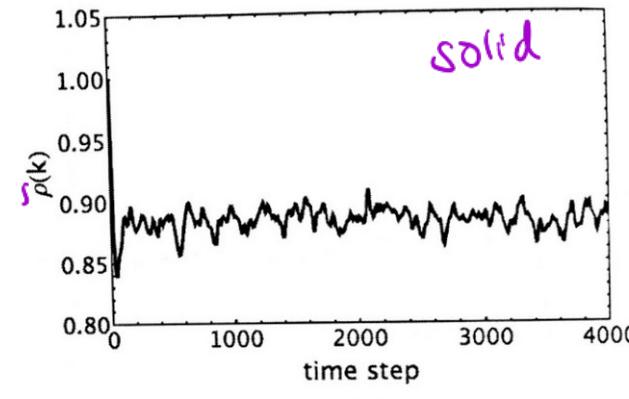
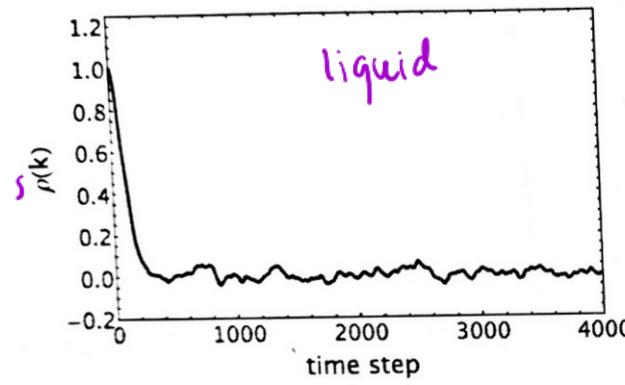


Figure 6.9 Order parameters for solid structures from Eq. (G.48). (a) $\rho(\mathbf{k})$ for the equilibration run shown in Figure 6.5. with $\rho^* = 0.55$. (b) $\rho(\mathbf{k})$ for the solid structure shown in Figure 6.8b with a density of $\rho^* = 1.0$.

$$\tilde{\rho}(\vec{k}) = \frac{1}{N} \sum_{i=1}^N \cos(\vec{k} \cdot \vec{r}_i)$$

(reciprocal lattice vector e.g., $\vec{k} = \frac{2\pi}{a} (-1, 1, -1)$ FCC)

An example: Lennard-Jonesium

Analysis of trajectories: Correlations of quantities

Spatial correlation functions: **translational order parameter**

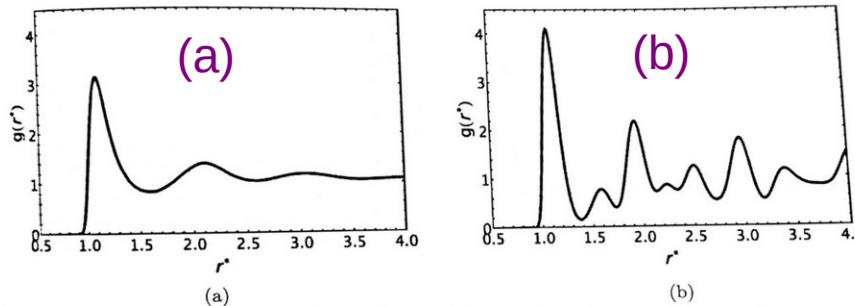


Figure 6.8 Radial distribution functions. (a) $g(r)$ for the equilibrated structure in Figure 6.6 with $\rho^* = 0.55$.
(b) Results from a similar calculation as in a, but with a density of $\rho^* = 1.0$.

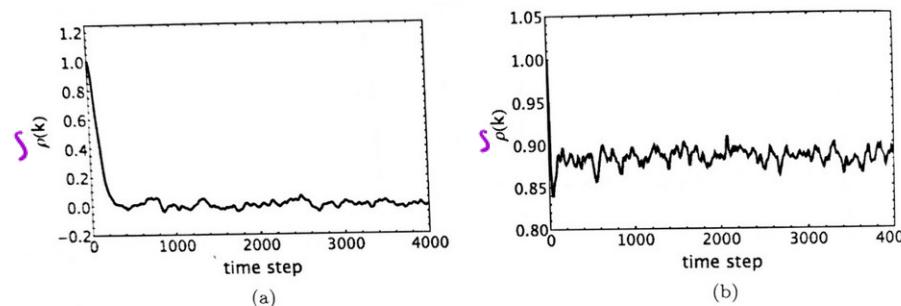


Figure 6.9 Order parameters for solid structures from Eq. (G.48). (a) $\rho(k)$ for the equilibration run shown in Figure 6.5, with $\rho^* = 0.55$. (b) $\rho(k)$ for the solid structure shown in Figure 6.8b with a density of $\rho^* = 1.0$.

As a complement
to $g(r)$

Which is a liquid? a)
Which is a solid? b)

An example: Lennard-Jonesium

Analysis of trajectories: Correlations of quantities

Time correlation functions: **velocity correlation function**

Time correlation function

$$c_{AB}(t) = \frac{\langle (A(t) - \langle A \rangle)(B(t' = 0) - \langle B \rangle) \rangle}{\sigma_A \sigma_B}$$

Autocorrelation function

$$c_{AA}(t) = \frac{\langle (A(t) - \langle A \rangle)(A(t' = 0) - \langle A \rangle) \rangle}{\sigma_A^2}$$

An example: Lennard-Jonesium

Analysis of trajectories: Correlations of quantities

Time correlation functions: **velocity correlation function**

Velocity autocorrelation function

$$c_{vv}(t) = \frac{\langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle}{\langle v^2 \rangle} = \frac{m}{3k_B T} \langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle$$

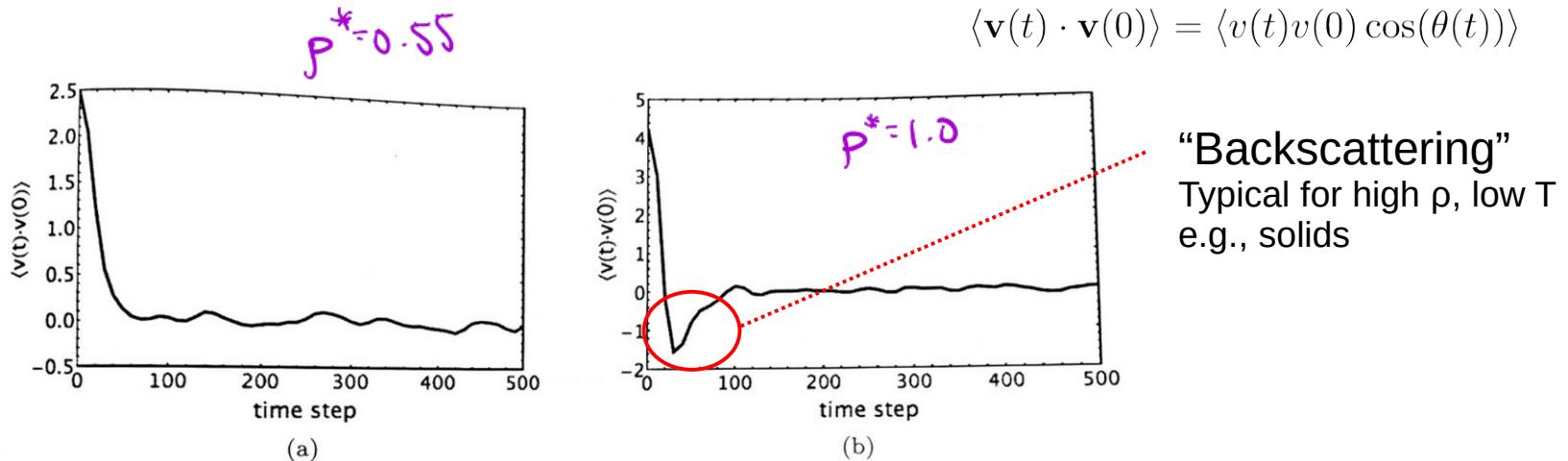


Figure 6.10 Velocity autocorrelation function from Eq. (6.28). Note that the statistics for these results are not particularly good. (a) $\langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle$ for the simulation results shown in Figure 6.6 with $\rho^* = 0.55$. (b) $\langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle$ for the solid structure shown in Figure 6.8b with a density of $\rho^* = 1.0$.

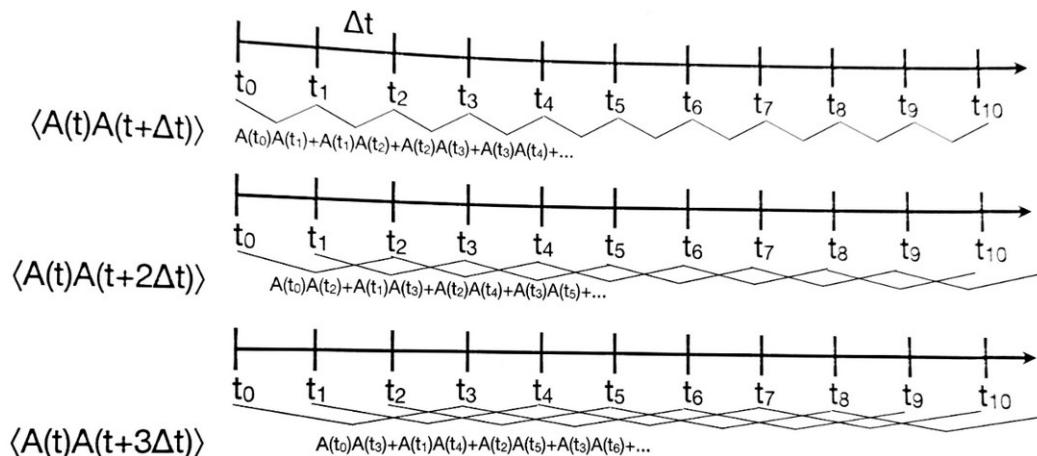
An example: Lennard-Jonesium

Analysis of trajectories: Correlations of quantities

Time correlation functions: **velocity correlation function**

Computing the velocity autocorrelation function
(for better statistics within a single trajectory)

vs average over N atoms (thermodynamic limit)
average over many trajectories



Use each point in MD simulation as starting point

Figure 6.16 Calculation of a time autocorrelation function showing how one set of time steps can be used to generate good averages using Eq. (6.38). At the top, averaging over each time pair $(t, t + \Delta t)$ yields $\langle A(t + \Delta t)A(t) \rangle$. In the middle, we accumulate all values to calculate $\langle A(t + 2\Delta t)A(t) \rangle$ and at the bottom $\langle A(t + 3\Delta t)A(t) \rangle$, and so on. Adapted from [172].

Examples: Diffusivity

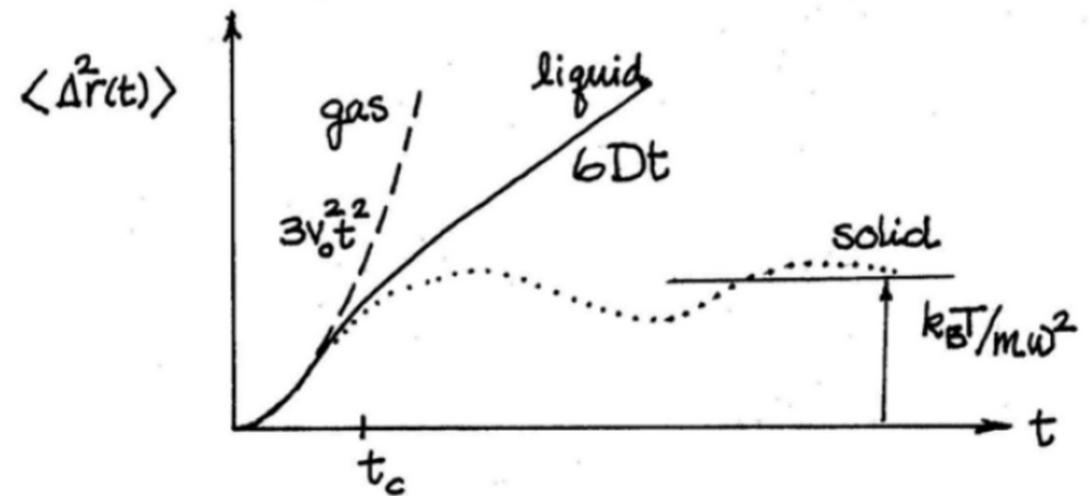
Analysis of trajectories: Correlations of quantities

Time correlation functions: diffusivity and mean square displacement

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

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

d is dimensionality



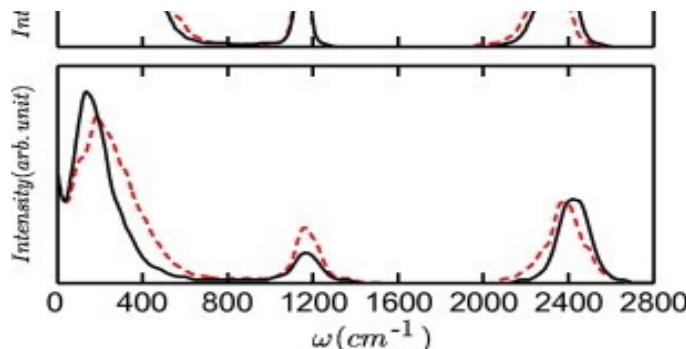
Examples: power spectrum of vibrations

Analysis of trajectories: Correlations of quantities

Time correlation functions: power spectrum for vibrational density of states

Wiener-Khintchine theorem: $\langle f(\tau) f(t + \tau) \rangle = \frac{1}{2\pi} \int \left| \int f(t) e^{-i\omega t} dt \right|^2 e^{i\omega t} d\omega$

Fourier Transform of velocity autocorrelation function is related to vibrational spectrum



Heavy water in bulk

Heavy water in hydration shell of supercritical ionic solutions

Mallik, Chandra, Chemical Physics 387(1):48-55

$$P(\omega) = m \int \langle \dot{\mathbf{r}}(\tau) \dot{\mathbf{r}}(t + \tau) \rangle e^{-i\omega t} dt$$

Phys. Chem. Chem. Phys., 2013, 15, 6608-6622

No calculation of phonon modes!

Basics: Thermodynamics (connecting to physical observables)

Basics: Laws of thermodynamics

Zeroth law: defines thermal equilibrium

If system **A** is in equilibrium with system **B** and system **B** is equilibrium with system **C**,
then system **A** is in equilibrium with **C**

First law: energy conservation

Total energy (into or out of; in the form of work, heat , matter) is conserved

Second law: arrow of time or irreversible processes

At equilibrium, the (overall total of) entropy **S** is maximized.

Corollary: Heat does not spontaneously flow from cold to hot

Third law: “freeze out”

Entropy approaches constant (often zero) value as $T \rightarrow 0\text{ K}$

Corollary: Only degrees of freedom that are excited at finite T contribute to the
thermodynamic properties

Basics: Stat Mech

Extensive variables — scale w system size
→ internal energy $E = U + K$
enthalpy $H = E + PV$

Thermodyn. potentials

Helmholtz $A = E - TS$

Gibbs $G = E - TS + PV$

Intensive variables — not scale w system size "natural variables"
e.g., T, P, μ

$$dG = \cancel{-SdT} + \cancel{VdP} + \sum_{i=1}^N \mu_i dn_i$$

$$dH = \cancel{TdS} + \cancel{VdP} + \sum_i \mu_i dn_i$$

$$dA = \cancel{-SdT} - \cancel{PdV} + \sum_i \mu_i dn_i$$

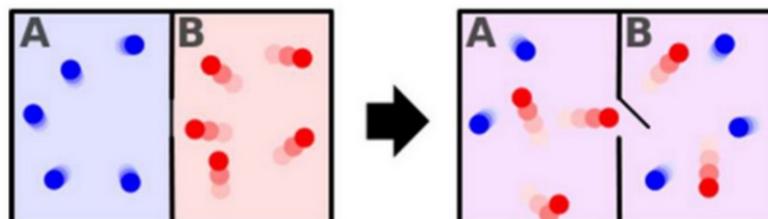
$$dE = \cancel{TdS} - \cancel{PdV} + \sum_i \mu_i dn_i$$

Basics: Entropy and Equilibrium

Entropy: a measure of the spontaneity or disorder of the system (classical thermo.)
a measure of number of microscopic configurations Ω (stat. mech definition)
a measure of informational value based on the
likelihood of communication events (information theory)

$$S = k_B \ln \Omega(N, V, E) = -k_B \sum_i P_i \ln P_i$$

$$\beta = \frac{1}{k_B T} = \frac{1}{k_B} \frac{dS}{dE} = \frac{d \ln \Omega}{dE}$$

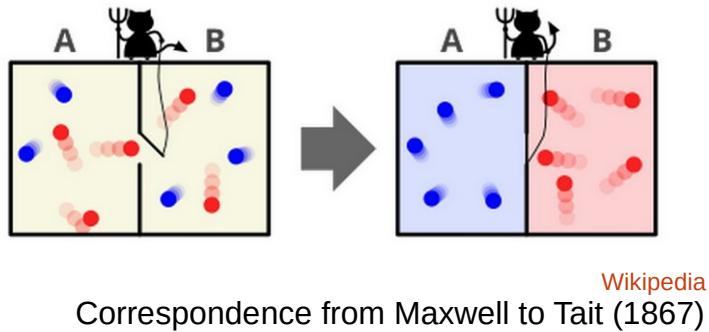


Any method involving the notion of entropy, the very existence of which depends on the second law of thermodynamics, will doubtless seem to many far-fetched, and may repel beginners as obscure and difficult of comprehension.

Willard Gibbs, Graphical Methods in the Thermodynamics of Fluids^[11]

Basics: Entropy and Equilibrium

Maxwell's demon:



... if we conceive of a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are as essentially finite as our own, would be able to do what is impossible to us. For we have seen that molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower molecules to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics.

Szilard and Brillouin (1920s): the act of measuring requires thermodynamic work
Total entropy is that of system and demon combined

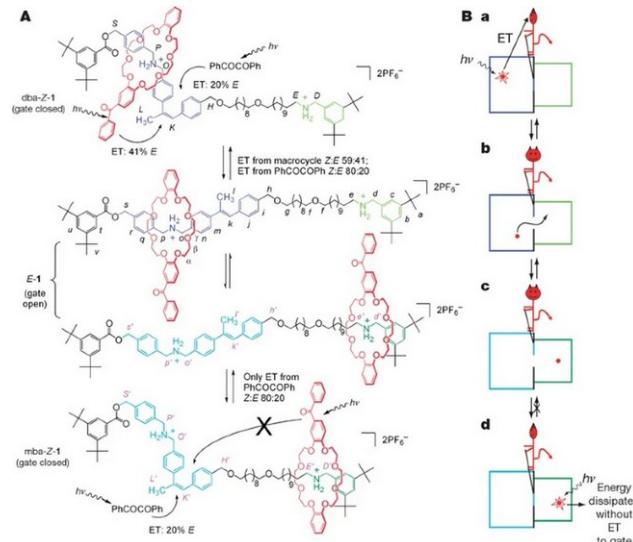
Landauer (1960): reversible thermodynamic processes need not expend work
however, the demon must eventually discard information,
which is an irreversible process that will increase entropy

Basics: Entropy and Equilibrium

Maxwell's demon: motor proteins and biological machines,
which drive chemical systems away from thermodynamic equilibrium
e.g., ATP-driven molecular motors for muscle contraction,
intracellular vesicle transport

Molecular Machine Ratchet

more probable for **ring molecule** to be on the left **binding site** than right **binding site**



A molecular information ratchet. Nature 445, 523–527 (2007).

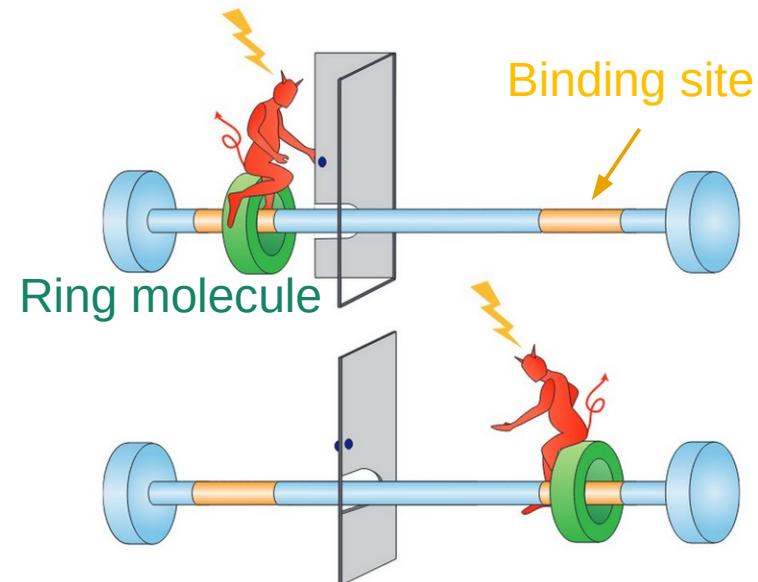


Figure 1 A schematic representation of the ‘information ratchet’ of Leigh and co-workers. The ring (green) can move along the axle (blue) provided that the gate (door) is open. The ring carries a mechanism for opening the gate (the demon) that is activated when energized by light, but this only works when the demon is close to the gate. The axle incorporates binding sites (orange) for the ring, and the site on the left is much closer to the gate. From the left-hand station the demon can open the gate, but from the right-hand station he cannot reach. The ring thus spends more time on the right, even though thermodynamics dictates that it should spend more time on the left.

Basics: Stat Mech

Phase space

N atoms $N \sim 10^{23}$
mechanical d.o.f.: 3N positions $\vec{r}_1, \vec{r}_2, \vec{r}_3, \dots, \vec{r}_N$ } "phase space"
 3N momenta $\vec{p}_1, \vec{p}_2, \dots, \vec{p}_N$

trajectory \rightarrow locus of points in motion
through phase space

connect microstates \rightarrow macrostate
e.g., time avg $E_{\text{obs}} = \langle E \rangle = \frac{1}{t_0} \int_0^{t_0} E(t') dt'$

What excitations accessible?



Basics: Stat Mech

The partition function

$$Z = \sum_i e^{-\beta E_i}$$

Boltzmann factor

$$\beta = \frac{1}{k_B T}$$

ensemble avg^s

$$\langle O \rangle = \frac{1}{Z} \sum_i O_i e^{-\beta E_i}$$
$$= \frac{\sum_i O_i e^{-\beta E_i}}{\sum e^{-\beta E_i}}$$

internal E

$$\langle E \rangle = \frac{1}{Z} \sum_i E_i e^{-\beta E_i}$$
$$= - \frac{\partial \ln(Z)}{\partial \beta}$$

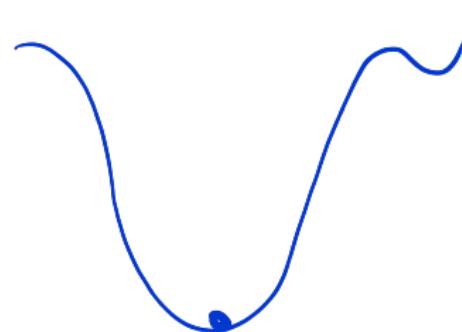
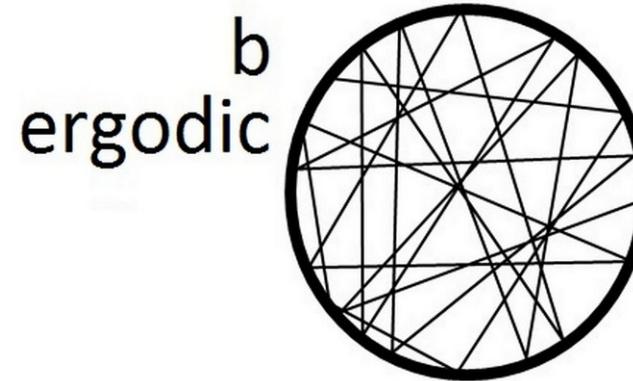
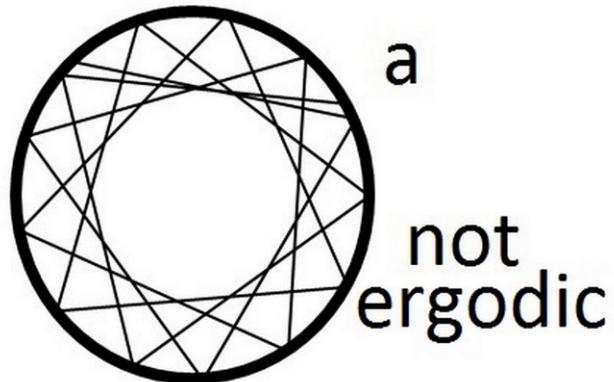
pressure

$$P = k_B T \left(\frac{\partial \ln Z}{\partial V} \right)_T$$

What are the excitations in the system?

Basics: Stat Mech

Ensembles and Ergodicity



Basics: Stat Mech

Ensembles and Ergodicity

Basics: Stat Mech

Temperature in stat mech

Equipartition theorem for energy:

Kinetic energy is shared equally among all degrees of freedom
temperature is related to the average kinetic energy of system

$$\langle KE \rangle = \left\langle \sum_{i=1}^N \frac{m_i \mathbf{v}_i^2}{2} \right\rangle = \frac{1}{2} D_f k_B T(t) = \frac{3}{2} N k_B T(t)$$

e.g., ideal gas of
 N particles and
3 translational d.o.f:
 v_x, v_y, v_z

Consequences:

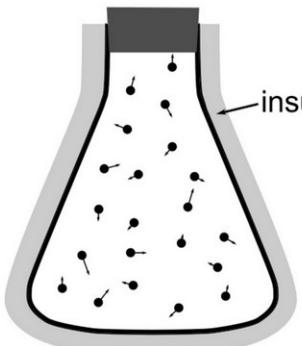
Heavy atoms move slower than light atoms at equilibrium temperature T

Macroscopic Temperature as a time-average of instantaneous temperature

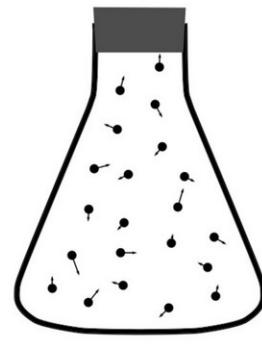
$$\langle T(t) \rangle = \frac{1}{\tau} \sum_i^\tau T(t_i)$$

Basics: Ensembles

Statistical ensembles



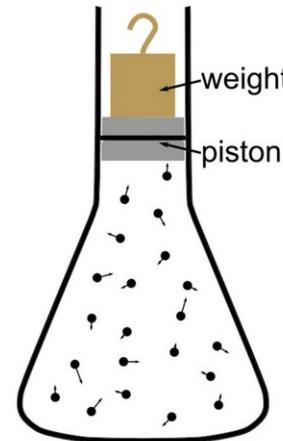
Microcanonical
(const. NVE)



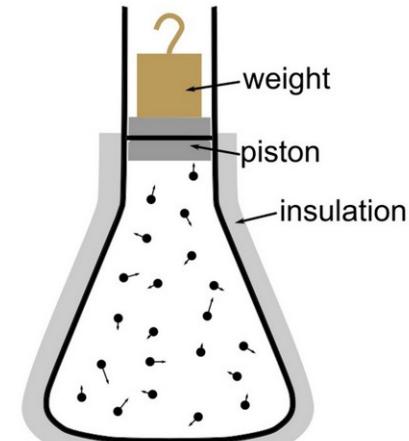
Canonical
(const. NVT)



Grand Canonical
(const. μVT)



Gibbs or
Isobaric-isothermal
(const. NPT)



Enthalpy or
Isoenthalpic-isobaric
(const. NPH) $H=E+PV$

Closed, isolated

Closed

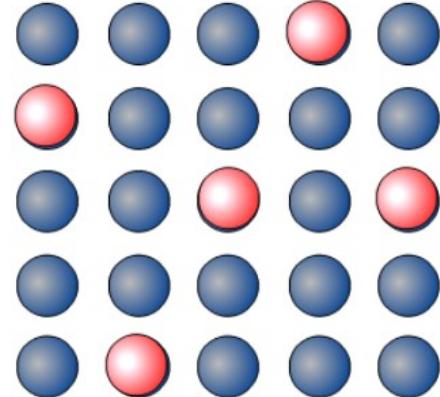
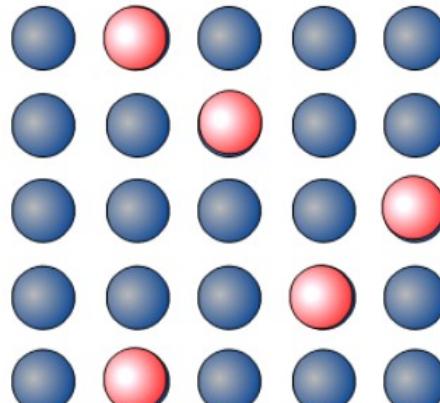
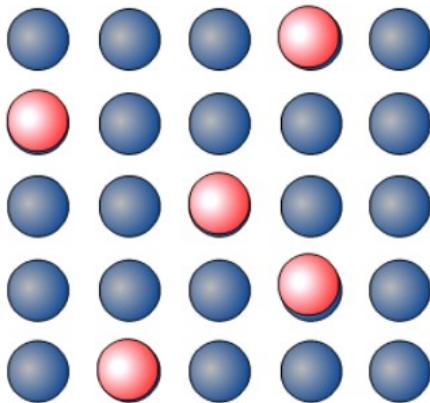
Open

Closest to “lab”
conditions

Nzjacobmartin - CC BY-SA 4.0.

Basics: Stat Mech

Microstate versus macrostate



Basics: Particle statistics and occupation

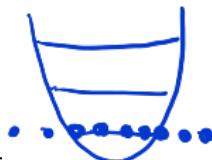
Bose-Einstein: number of configurations to distribute N_i indistinguishable particles in A_i energy levels;
no limit on occupancy of state i ; e.g., phonons, photons (particles with integer spin)

Count
microstates

$$W = \prod_i \frac{(N_i + A_i)!}{N_i! A_i!}$$

Extent of
occupation

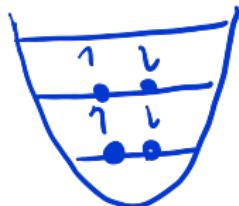
$$\frac{N_i}{A_i} = \frac{1}{B \exp(\epsilon_i/k_B T) - 1}$$



Fermi-Dirac: number of configurations to distribute N_i indistinguishable particles in A_i energy levels;

each unique state has occupancy 0 or 1 (Pauli exclusion principle);

e.g., electrons, nuclei, protons (particles with spin in odd half integer values, e.g., 1/2, 3/2, 5/2)



$$W = \prod_i \frac{A_i!}{N_i!(A_i - N_i)!}$$

$$\frac{N_i}{A_i} = \frac{1}{B \exp(\epsilon_i/k_B T) + 1} = \frac{1}{\exp((\epsilon_i - \mu)/k_B T) + 1}$$

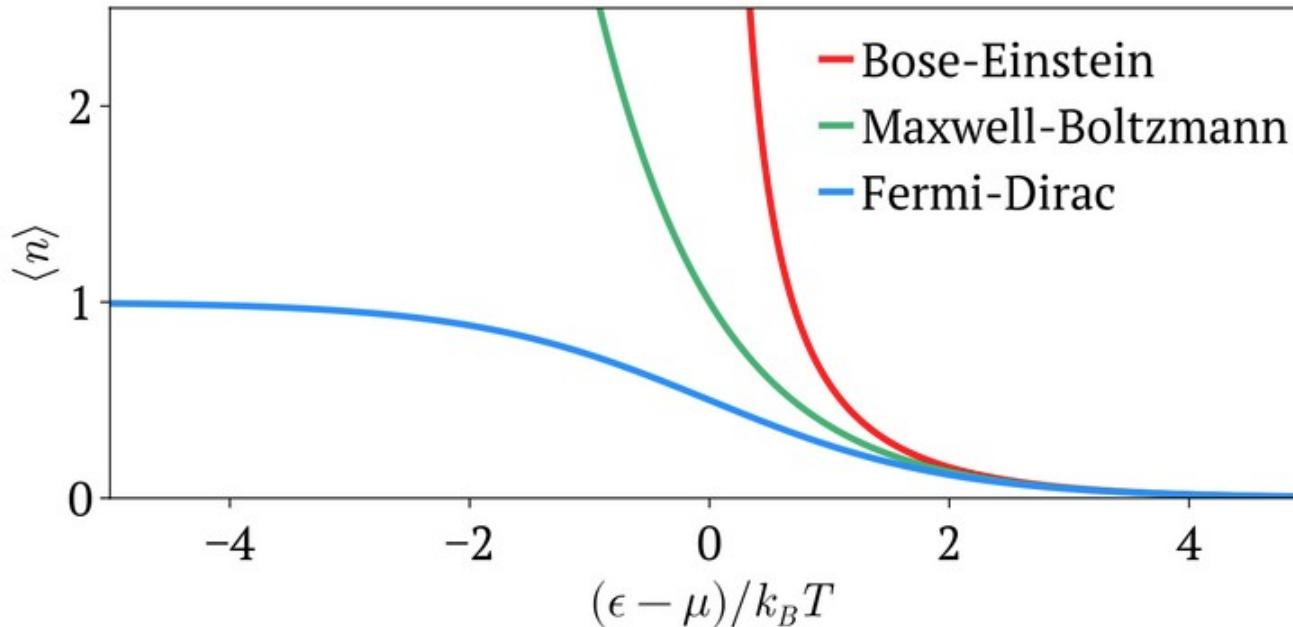
Probability of occupation

Maxwell-Boltzmann: distinguishable particles

limit of B-E and F-D distributions at temperatures \gg excitation energies
easier to work with mathematically

$$\frac{N_i}{A_i} = \frac{1}{B \exp(\epsilon_i/k_B T)}$$

Basics: Particle statistics and occupation



Basics: Ensembles

Microcanonical Ensemble, NVE

microstate : set of $\{n_i\}$ occupations

$$N = \sum_i n_i$$

energy: $E = \sum_j n_j \epsilon_j$

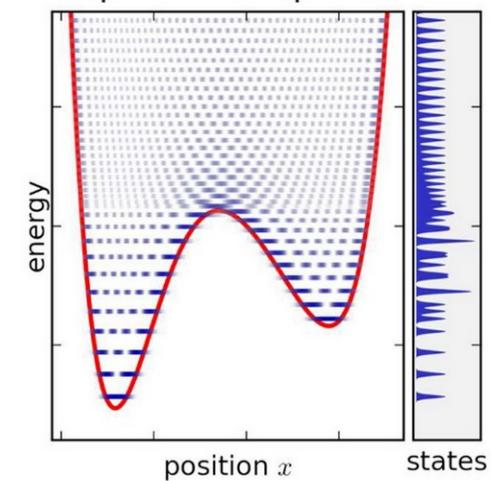
$$\langle E \rangle = \frac{1}{N_{\text{config}}} \sum_{\alpha} n_{\alpha} E_{\alpha} = \sum_{\alpha}^{N_{\text{config}}} p_{\alpha} E_{\alpha}$$

\uparrow
config in α

$$= \frac{n_{\alpha}}{N_{\text{config}}}$$

$$p_{\alpha} = \frac{\omega_{\alpha}}{\sum_{\alpha} \omega_{\alpha}}$$

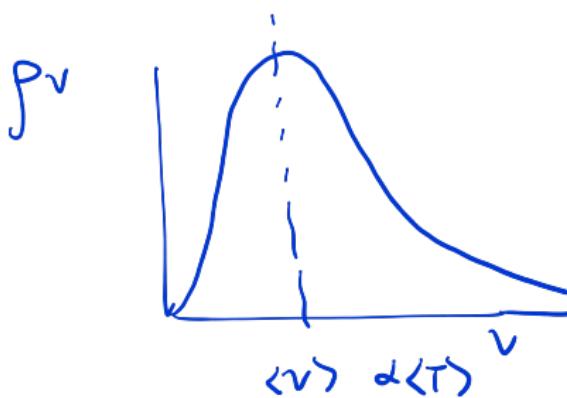
$$\omega_{\alpha}^{\text{NVE}} = \delta(E(\bar{r}^n, \bar{p}^n) - E_{\alpha})$$



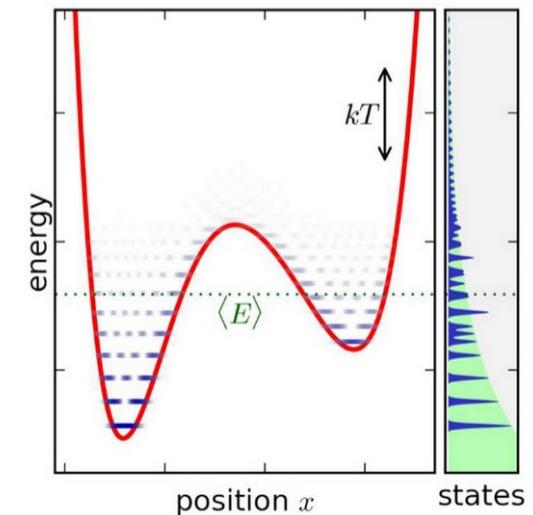
Basics: Ensembles

Canonical Ensemble, NVT

weighting function: $w_{\alpha}^{NVT} = e^{-E_{\alpha}/k_B T}$ $\approx e^{-E_{\alpha}/k_B T}$ $\sim 0.025 \text{ eV}$
 25 meV



$$\langle E \rangle = \frac{1}{\sum_{\alpha} e^{-E_{\alpha}/k_B T}} \sum_{\alpha} E_{\alpha} e^{-E_{\alpha}/k_B T}$$



Basics: Ensembles

Thermodynamic ensembles^[17]

	Microcanonical	Canonical	Grand canonical
Fixed variables	E, N, V	T, N, V	T, μ, V
Microscopic features	Number of microstates	Canonical partition function	Grand partition function
	W	$Z = \sum_k e^{-E_k/k_B T}$	$\mathcal{Z} = \sum_k e^{-(E_k - \mu N_k)/k_B T}$
Macroscopic function	Boltzmann entropy	Helmholtz free energy	Grand potential
	$S = k_B \log W$	$F = -k_B T \log Z$	$\Omega = -k_B T \log \mathcal{Z}$

Wikipedia

NVE: maximize entropy S at equilibrium

NVT: minimize Helmholtz free energy F at equilibrium

Basics: Stat Mech

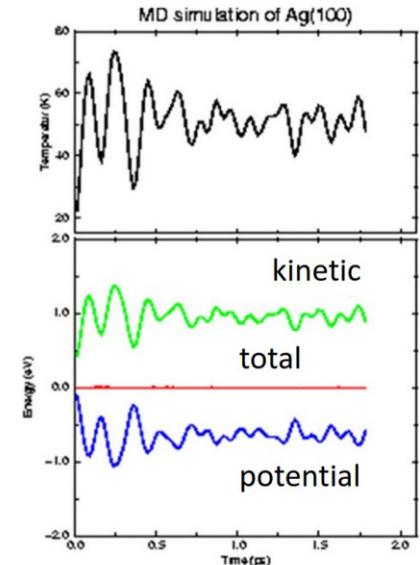
Temperature fluctuations, e.g., in NVE ensemble

Total energy is fixed while kinetic energy (and T) fluctuate by equal and opposite values

(Fluctuations in total energy will happen practically due to finite timestep in integration
but should be small relative to the fluctuations in K)

$$\langle \sigma_T^2 \rangle = \langle T^2 \rangle - \langle T \rangle^2 = \frac{k_B T^2}{N C_V}$$

For large systems, fluctuations in T is negligible



Basics: Stat Mech

Energy fluctuations, e.g., in NVT ensemble

Temperature is fixed while total energy is allowed to fluctuate

$$\langle E \rangle = -\frac{\partial \ln \mathcal{Z}}{\partial \beta} \quad \langle E^m \rangle = \frac{1}{\mathcal{Z}} \sum_j E_j^m e^{-\beta E_j}$$

Can show that:

$$\langle (E - \langle E \rangle)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2 = \sigma_E^2 = \frac{\partial^2 \ln \mathcal{Z}}{\partial \beta^2}$$

Basics: Ensembles

Equivalence of ensembles for the thermodynamic limit

Integrating Newton's equations with MD naturally corresponds to NVE ensemble

But other ensembles such as NVT are more convenient
(e.g., better correspondence with experimental conditions)

Is there still valid dynamical information from these other ensembles?

Yes, up to a point, but practically always dealing with finite sized systems

e.g., NVT approaches NVE in thermodynamic limit (fluctuations in E vanish)

$$\frac{\sigma_E}{\langle E \rangle} = \frac{\sqrt{\langle E^2 \rangle - \langle E \rangle^2}}{\langle E \rangle} \sim \frac{\sqrt{N}}{N} \rightarrow 0 \quad \text{for } N \rightarrow \infty$$

Control of system variables (T, P)

Strong-coupling methods: scale variable to exact derived values
e.g., velocity rescaling

Weak-coupling methods: scale variable towards derived values
e.g., Berendsen thermostat

Stochastic methods: constrain variable to a pre-determined distribution function
e.g., Andersen thermostat, Langevin thermostat

Extended system dynamics: add degrees of free to include T, P
e.g., Nosé-Hoover thermostat

↓
Inc. implementation complexity



Frenkel & Smit on some of these less Newtonian methods:
“these algorithms cannot be derived from the basic laws of mechanics,
their choice is limited by the creativity of their authors”

We discuss a few examples of thermostats

Control of system variables (T, P)

Velocity re-scaling $T \propto v^2$

Instantaneous temperature

$$T(t) = \frac{2K(t)}{3Nk_B}$$

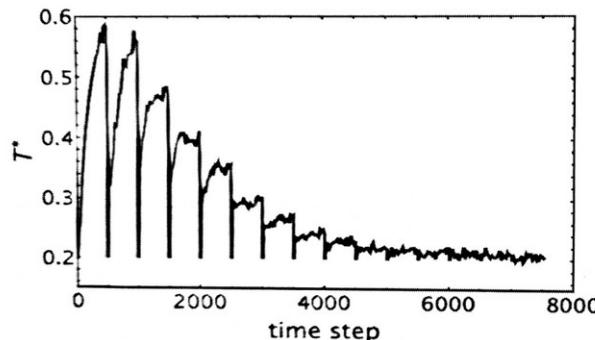


Figure 6.11 Variation of T^* in a velocity rescaling simulation.

Scale velocities to get to desired temperature T_{set}

$$v_{i,\alpha}^{new} = \sqrt{\frac{T_{set}}{T(t)}} v_{i,\alpha}$$

$$\equiv \lambda$$

Cons: Affects natural dynamics of the system
Not a canonical distribution

Other variations:

“Canonical sampling through velocity rescaling.” J. Chem. Phys. 126, 014101 (2007)

Control of system variables (T, P)

Berendsen thermostat (weakly coupled variation of velocity rescaling)

Idea: Couple to an external bath at T_{set} that supplies or removes thermal energy as needed

Set velocities at each step such that

$$\frac{dT(t)}{dt} = \frac{1}{\tau} (T_{set} - T(t))$$

Strength of coupling to bath via τ

Choose τ carefully, typically $\tau = 0.1 - 0.4$ ps

Small τ : too aggressive coupling, Large τ : too slow coupling

Scale velocities by: $\lambda = \sqrt{1 + \frac{\delta t}{\tau} \left(\frac{T_{set}}{T(t)} - 1 \right)}$

Simplifies to velocity re-scaling
for $\tau = \delta t$ timestep

Cons: No guarantee of sampling canonical distribution

See also Bussi-Donadio-Parrinello thermostat:

extension of Berendsen thermostat; allows for proper sampling of canonical ensemble

Rescaling of velocities is more gentle:

done to kinetic energy that is stochastically chosen from canonical distribution for K

Control of system variables (T, P)

Andersen thermostat (stochastic coupling)

Idea: influence the temperature by resetting velocity of random particle via “collision”

Probability of collision with collision frequency ν $\mathcal{P}(t) = \nu \exp(-\nu t)$
Assumed to follow Poisson distribution

1. Start from initial configuration, integrate equations of motion for δt
2. m atoms are randomly selected to undergo a collision
3. For each m particles, a new velocity from Maxwell-Boltzmann distribution corresponding to T_{set} is chosen

Pros: Yields a canonical distribution

Cons: Does not conserve momentum or energy

This will affect transport and non-equilibrium properties:
rates, diffusion, thermal conductivity, viscosity

Ok for static properties

Control of system variables (T, P)

Nosé-Hoover thermostat (extended system dynamics)

Idea: introduce a heat bath as an integral part of the system,
i.e., as an additional degree of freedom s

Modify the Hamiltonian to include a fictitious coordinate s
with an effective “mass” Q and velocity ds/dt

$$\mathcal{H} = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i s^2} + \frac{1}{2} \sum_{i,j \neq i} U(\mathbf{r}_j - \mathbf{r}_i) + \frac{\mathbf{p}_s^2}{2Q} + L \frac{\ln s}{\beta} \equiv \frac{1}{k_B T}$$

The extra degree of freedom s represents a friction that slows down or
accelerates particles until the temperature reaches T_{set}

Temperature is not fixed but *tends* to a target value

$$L = \text{degrees of freedom} = 3N + 1$$

Control of system variables (T, P)

Nosé-Hoover thermostat (extended system dynamics)

Because the bath is part of the system, it too has equations of motion

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{f}_i - \zeta m_i \mathbf{v}_i \quad \zeta = \frac{\partial \ln s}{\partial t}$$

$$\frac{d\zeta(t)}{dt} = \frac{1}{Q} \left[\sum_{i=1}^N m_i \frac{\mathbf{v}_i^2}{2} - \frac{L}{2\beta} \right]$$

Can modify the velocity Verlet algorithm to implement equations of motion for NH thermostat

$$\begin{aligned}\mathbf{r}_i(t + \delta t) &= \mathbf{r}_i(t) + \mathbf{v}_i(t)\delta t + \left(\frac{\mathbf{f}_i(t)}{m_i} - \zeta(t)\mathbf{v}_i(t) \right) \frac{\delta t^2}{2} \\ \mathbf{v}_i(t + \underline{\delta t/2}) &= \mathbf{v}_i(t) + \frac{\delta t}{2} \left(\frac{\mathbf{f}_i(t)}{m_i} - \zeta(t)\mathbf{v}_i(t) \right) \\ \mathbf{f}_i(t + \delta t) &= \mathbf{f}_i(\mathbf{r}_i(t + \delta t)) \\ \underline{\zeta(t + \delta t/2)} &= \zeta(t) + \frac{\delta t}{2Q} \left[\sum_{i=1}^N m_i \frac{\mathbf{v}_i(t)^2}{2} - \frac{3N+1}{2} k_B T \right].\end{aligned}$$

Final step:

$$\begin{aligned}\mathbf{v}_i(t + \delta t) &= \frac{\left[\mathbf{v}_i(t + \delta t/2) + \frac{\delta t}{2} \frac{\mathbf{f}_i(t+\delta t)}{m_i} \right]}{1 + \frac{\delta t}{2} \zeta(t + \delta t)} \\ \zeta(t + \delta t) &= \zeta(t + \delta t/2) + \frac{\delta t}{2Q} \left[\sum_{i=1}^N m_i \frac{\mathbf{v}_i(t + \delta t/2)^2}{2} - \frac{3N+1}{2} k_B T \right]\end{aligned}$$

Control of system variables (T, P)

Nosé-Hoover thermostat (extended system dynamics)

Pros:

- Sampling canonical statistics for real system
(microcanonical statistics for combined system of bath and real system)
- *Mild* thermostat (affects all particle motions a bit but none strongly);
dynamics approaches Newtonian dynamics
- Conserves *virtual* momentum and a Hamiltonian-like quantity
(useful for checking errors in code)

Cons:

- Does not preserve area of phase space of the real coordinates and momenta
- Not always ergodic

See the works of Tuckerman and Martyna

Control of system variables (T, P)

“The Flying Ice Cube effect”

Gradual bleeding of KE from high-frequency motions to low-frequency motions

e.g., bond stretching modes → angle bending modes

e.g., vibrational, intramolecular d.o.f. → translational kinetic energy

In violation of the equipartition theorem

Effects of repeated velocity rescaling

Example:

conformations of circular DNA coupled to
Berendsen thermostat

Too frequent velocity rescaling leads to
fictitious open circle conformations

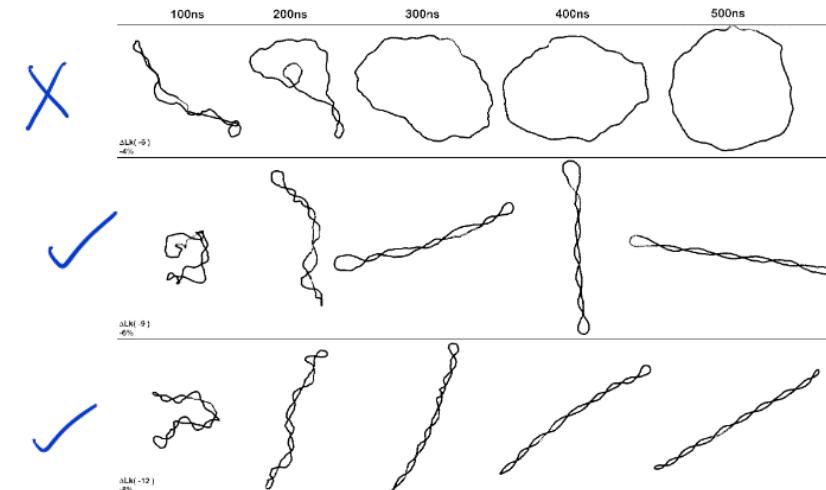


FIGURE 5. Evolution of the structure of model supercoiled DNAs in long MD simulations with coupling to a constant temperature bath. The model has 1500 bp. The expected writhe is about 70% of the linking deficit. This is indeed the case after a long time for two cases ($\Delta Lk = -9$ and $\Delta Lk = -12$), but the rotational motion of the molecule with $\Delta Lk = -4$ causes it to open into a circular form. After several hundred nanoseconds, almost all the kinetic energy in all three models is associated with the translation of the center of mass and rotational motion about the center of mass.

Enhanced sampling methods

Sampling rare events by biasing the potential

bias to sample transition state

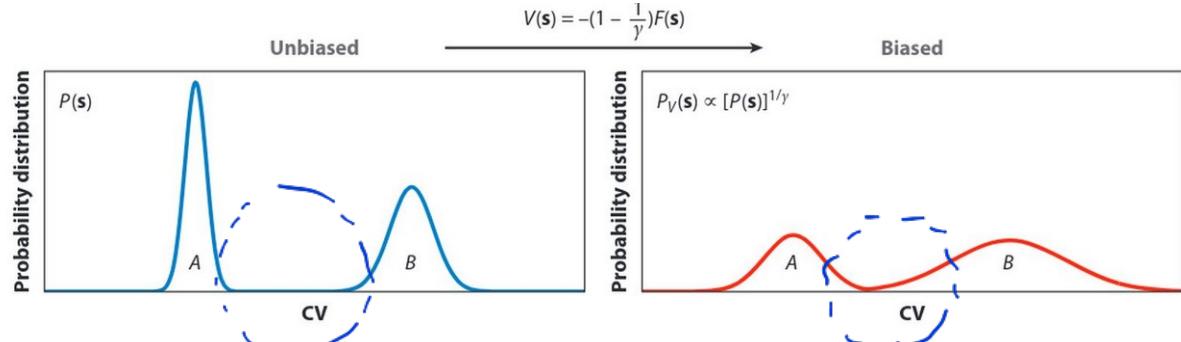
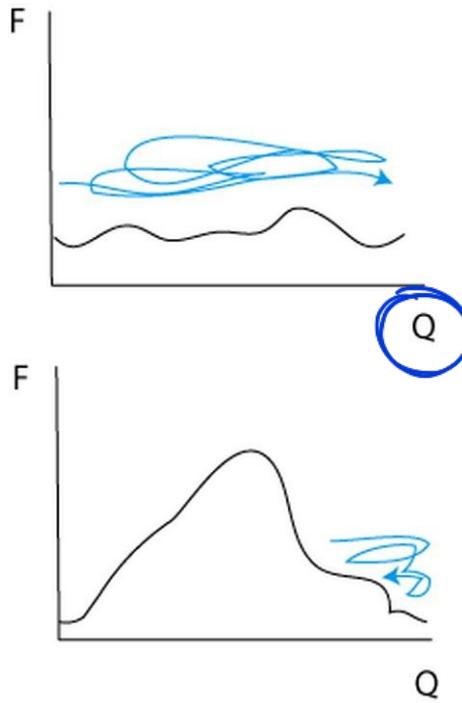


Figure 2

Model probability distribution of a system with two metastable states A and B separated by a transition region that is rarely visited. The effect of the bias in Equation 7 is to broaden the two peaks [i.e., enhance collective variable (CV) fluctuations] such that the transition region is more easily visited and transitions between the two states are more likely to occur.

In metadynamics, the dynamical evolution of the system is altered by the addition of an external bias potential that is periodically updated via the following stochastic iteration:

$$V_n(\mathbf{s}) = V_{n-1}(\mathbf{s}) + G(\mathbf{s}, \mathbf{s}_n) \exp\left[-\frac{1}{\gamma-1} \beta V_{n-1}(\mathbf{s}_n)\right], \quad (13)$$

Annual Rev. Phys. Chem. 67, 159-184 (2016)

$Q \equiv$

Enhanced sampling methods

(Collective variables:
coarse-grained order parameter that distinguishes between relevant metastable states

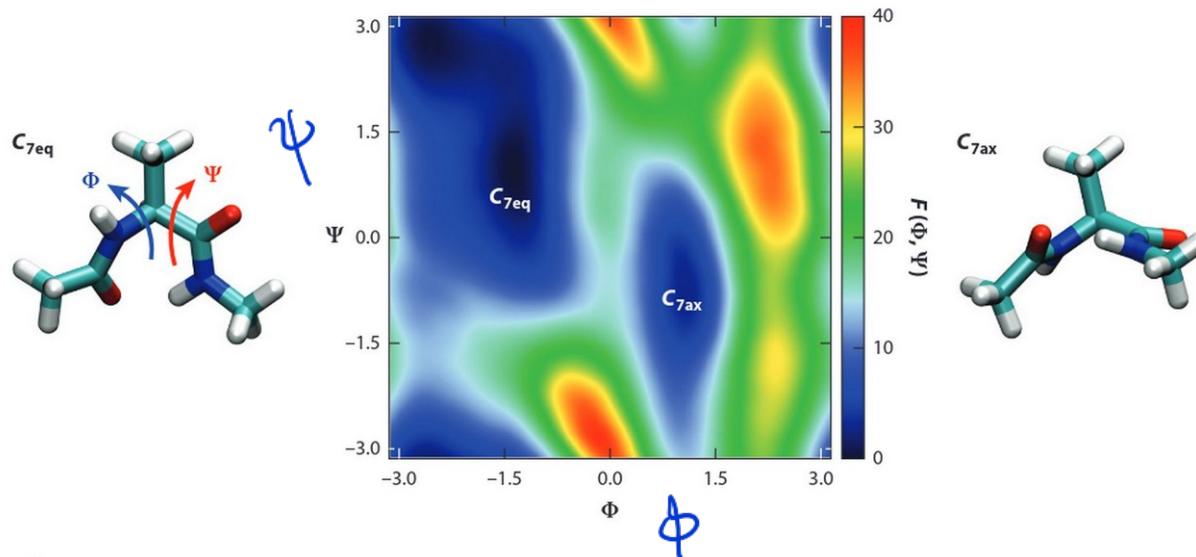


Figure 1

Free energy surface (FES) $F(\Phi, \Psi)$ for alanine dipeptide in vacuum in terms of its two backbone dihedral angles Φ and Ψ . $F(\Phi, \Psi)$ is in units of $k_B T$, and Φ and Ψ are given in radians. The FES is characterized by two metastable states, $C_{7\text{eq}}$ and $C_{7\text{ax}}$, separated by a free energy barrier of approximately $14 k_B T$, making transitions between them very rare.

Enhanced sampling methods

Guidelines for choosing collective variables

- Clearly distinguish between initial state, final state, and intermediates
- Describe all relevant timescales to the process of interest, particularly slow modes
- Limit the number; otherwise, it will take a while to fill the free energy surface

Usually requires some initial understanding of the problem or system at hand

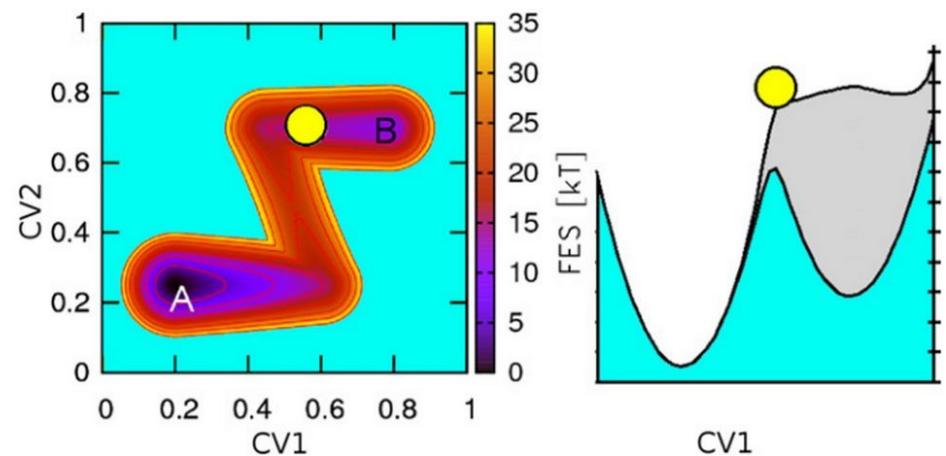
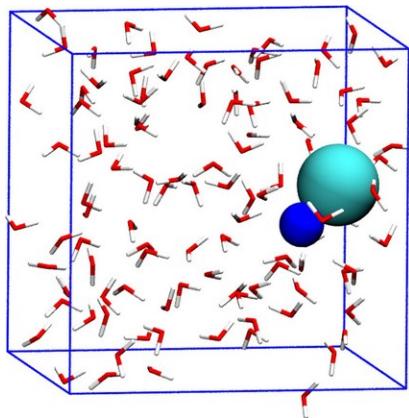


Figure 2. The effect of neglecting a relevant degree of freedom. Left side: 2D Z-shaped potential. Right side: the trajectory of a metadynamics simulation generated using only s_1 as CV. Transitions from A to B are properly described by CV1, causing strong hysteresis in the reconstructed free energy.

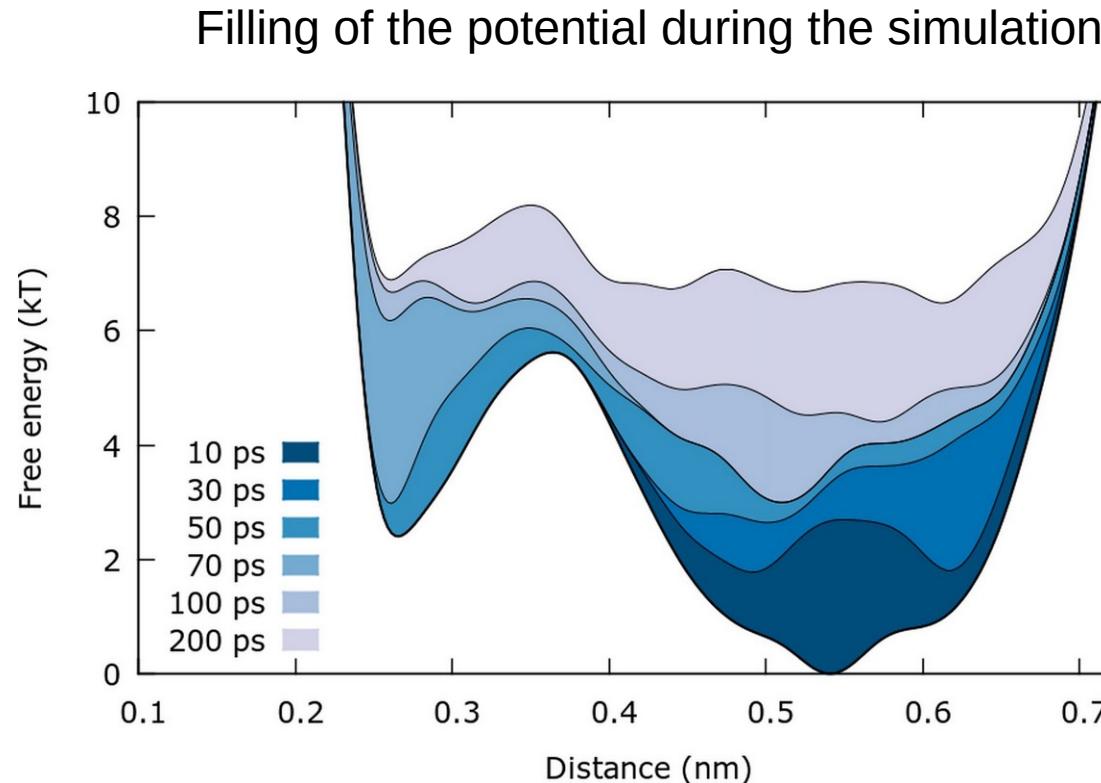
Enhanced sampling methods

An example: association/dissociation of NaCl in aqueous solution



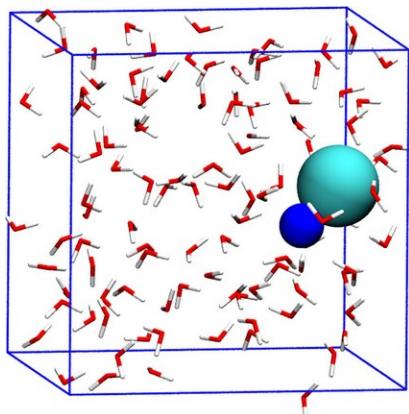
CV = distance between Na^+ and Cl^-

How does the surrounding aqueous environment (collective motions of solvent) affect the dissociation energy?



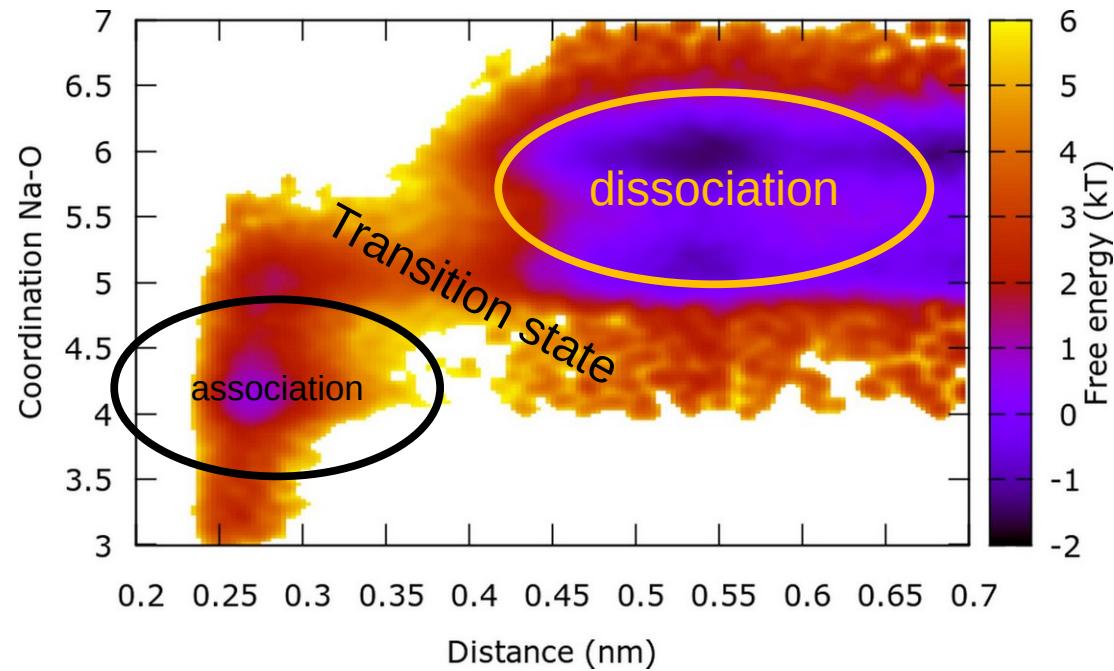
Enhanced sampling methods

An example: association/dissociation of NaCl in aqueous solution



Adding the coordination of Na-O as a CV for the solvent reveals the process of dissociation/association related to coordination environment with solvent

Free energy landscape and the role of the solvent



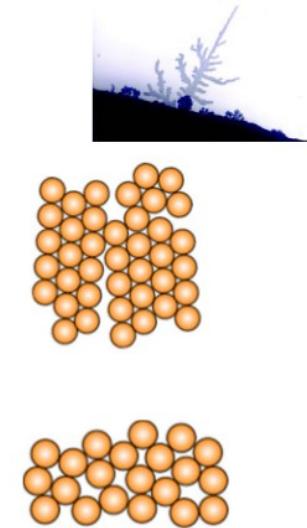
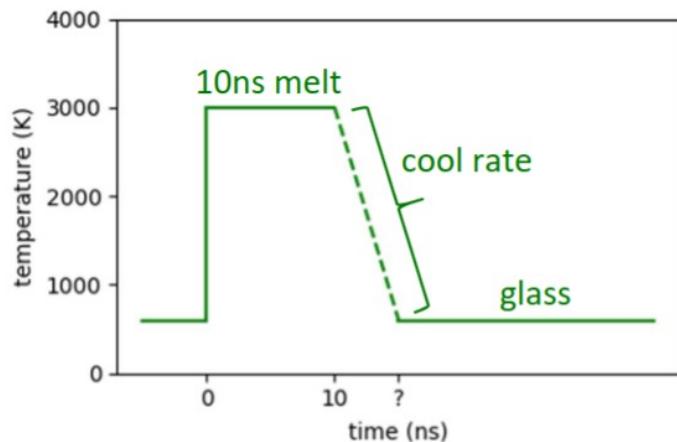
Applications of MD: Phase diagrams

Making glass electrolytes

Dendrites preferentially form along grain boundaries

Glasses can be fabricated without grain boundaries

MD preparation: crystal → melt → quench



Applications of MD: Phase diagrams

Li content and structure control conductivity

Optimization recipe

1. Reduce Li content
2. Avoid ordered states

Disorder is good for transport!

