Enhanced molecular dynamics

Ensembles

- micro-canonical ensemble NVE
- canonical ensemble NVT
- isothermal-isobaric NPT
- grand-canonical μVT
- isobaric-isoenthalpic NPH
- non-equilibrium

Temperature in MD

• In canonical (NVT) ensemble

$$\left(\sum_{I=1}^{N} \frac{\mathbf{P}_{I}^{2}}{2M_{I}}\right)_{\mathsf{NVT}} = \frac{3}{2}Nk_{B}T$$

• Maxwell-Boltzmann distribution of velocities ($\alpha = x, y, z$)

$$P(v_{\alpha}) = \sqrt{\frac{M}{2\pi k_B T}} \exp\left(-\frac{M v_{\alpha}^2}{2k_B T}\right)$$

• Introduce *instantaneous* temperature

$$\mathcal{T}\left(\mathbf{P}^{N}\right) = rac{1}{3Nk_{B}}\sum_{I=1}^{N}M_{I}v_{I}^{2}$$

• \mathcal{T} fluctuates in time (in finite systems), the average gives the temperature of microcanonical system

$$T_N = \frac{1}{M} \sum_{m=1}^M \mathcal{T}(t_m)$$

Ways to control temperature

One might want a constant temperature or a heating/cooling of the sample

- Rescaling of velocities
 - frequently scale the velocity by the ratio between the wanted and the instantaneous (or recent) temperature
 - does not lead to energy conservation or physical trajectories
 - very good in initial stages of equilibrating the system (*e. g.* after assigning random positions to the particles)
- Simulated annealing/heating
 - Scale atomic velocities at each time step; easy when using velocity Verlet
- Thermostats ("heat bath")
 - Andersen thermostat
 - Nosé-Hoover thermostat/thermostat chains

Nosé-Hoover thermostat

Attach a "reservoir" to the system which is able to adsorb and depose heat

$$\ddot{\mathbf{R}}_{I} = \frac{\mathbf{F}_{I}}{M_{I}} - \zeta \dot{\mathbf{R}}_{I} ,$$
$$\dot{\zeta} = \frac{1}{Q} \left[\sum_{I=1}^{N} M_{I} \dot{\mathbf{R}}_{I}^{2} - 3N k_{b} T_{\text{target}} \right]$$

• If . . .

- $T > T_{target}$: Increase of $\dot{\zeta}$, (eventually) larger friction
- $T < T_{\text{target}}$: Decrease of $\dot{\zeta}$, (eventually) smaller friction
- Q is an effective mass (inertia) for the friction
 - determines the response time of the thermostat to deviations of the actual temperature from the target ${\cal T}$
 - couples to the frequencies of the system; usually selected as a characteristic frequency ω_n in the spectrum
- new constant of motion:

$$E_{\text{conserved}}^{\text{NVT}} = E_{\text{k,I}}\left(\left\{\dot{\mathbf{R}}_{I}\right\}\right) + E_{\text{pot}}\left(\left\{\mathbf{R}\right\}\right) + \frac{Q}{2}\zeta^{2} + 3Nk_{B}T\int_{t}\zeta\,dt$$

even though the dynamics is non-Hamiltonian

Nosé-Hoover thermostat chains

- standard Nosé-Hoover thermostat:
 - does not always yield ergodic dynamics e.g. harmonic oscillator
 - slow response time
- Nosé-Hoover thermostat chains: Thermostatting the thermostat(s)

$$M_{I}\ddot{\mathbf{R}}_{I} = \mathbf{F}_{I} - M_{I}\dot{\xi}\dot{\mathbf{R}}_{I}$$

$$Q_{1}^{n}\ddot{\xi}_{1} = \left[\sum_{I=1}^{N} M_{I}\dot{\mathbf{R}}_{I}^{2} - 3Nk_{b}T\right] - Q_{1}^{n}\dot{\xi}_{1}\dot{\xi}_{2}$$

$$Q_{k}^{n}\ddot{\zeta}_{k} = \left[Q_{k-1}^{n}\dot{\xi}_{k-1}^{2} - k_{b}T\right] - Q_{k}^{n}\dot{\xi}_{k}\dot{\xi}_{k+1}\left(1 - \delta_{kK}\right), \quad k = 2, \dots, K$$

$$Q_{1}^{n} = \frac{3Nk_{B}T}{\omega_{n}^{2}}, \quad Q_{k}^{n} = \frac{k_{B}T}{\omega_{n}^{2}}$$

• new energy, but is still conserved:

$$E_{\text{conserved}}^{\text{NVT}} = E_{\text{k,I}}\left(\{\dot{\mathbf{R}}_{I}\}\right) + E_{\text{pot}}\left(\{vR\}\right) + \sum_{k=1}^{N} \frac{Q_{k}^{n}}{2} \dot{\xi}_{k}^{2} + 3Nk_{B}T\xi_{1} + \sum_{k=2}^{N} k_{B}T\xi_{k}$$

Nosé-Hoover thermostat chains Observations

- ergodicity problem *e. g.* in harmonic oscillator solved
- much more efficient in imposing the target temperature
- trajectories differ from the physical ones, so does the conserved energy, but thermal averages are fine

Other ensembles

- constant pressure
 - Parrinello-Rahman dynamics for the cell axis
 - * similar to Nosé-Hoover thermostat; barostat
 - * one has to be very careful with the change of the basis set when changing the cell; there are ways to treat this within CPMD, please check manual
- Constant number of particles

Tricks in MD

- periodic boundary conditions
- simulated annealing
- multiple time scales / RESPA (not in CPMD)
- Ewald summation; efficiency/speed vs accuracy (not in CPMD)
- cell lists *etc* (not in CPMD)

Reactions, rare events

Constraints

One might want a constraint...

- as a reaction coordinate
- to prevent a reaction/change in system
- to increase time step (*e. g.* —CH₃ group)

Example: a fixed inter-atomic distance:

$$\sigma_1 \left(\mathbf{R}_I, \mathbf{R}_J \right) = \left(|\mathbf{R}_I - \mathbf{R}_J| - d_{IJ} \right)^2$$

Constraints Lagrangean formulation

• Lagrangean (multipliers λ):

$$\mathcal{L}' = \mathcal{L} - \sum_{lpha} \lambda_{lpha} \sigma_{lpha} \left(\mathbf{R}^N
ight)$$

• Equation of motion:

$$\frac{\partial}{\partial t} \frac{\partial \mathcal{L}'}{\partial \mathbf{P}_I} = \frac{\partial \mathcal{L}'}{\partial \mathbf{R}_I}$$
$$M_I \ddot{\mathbf{R}}_I = -\frac{\partial E}{\partial \mathbf{R}_I} - \sum_{\alpha} \lambda_{\alpha} \frac{\partial \sigma_{\alpha}}{\partial \mathbf{R}_I}$$
$$= \mathbf{F}_I + \sum_{\alpha} \mathbf{G}_I (\alpha)$$

• λ_{α} (in principle) solved by requiring

$$\frac{\partial \dot{\sigma}_{\alpha}}{\partial t} = \frac{\partial \dot{\mathbf{P}} \cdot \nabla \sigma_{\alpha}}{\partial t}$$

• SHAKE/RATTLE algorithm: Constraints solved iteratively, one by one

Restraints

- Similar to constraints, but harmonic and allows for deviations
- For example distance, $E_{\text{restraint}} = k_r (|\mathbf{R}_I \mathbf{R}_J| d_0)^2$, $d_0 = \text{target distance}$, $k_r = \text{force constant of restraint}$
- Can be either attractive $(k_r > 0)$ or repulsive $(k_r < 0; e. g. in angles)$
- Does not in general allow a reduction of the time step (rather would increase it), but can be used *e.g.* to hold atoms in a molecule together
- Note: Energy no longer conserved

Constraint types

- Distance: $|\mathbf{R}_I \mathbf{R}_J| d_0$, $d_0 = \text{target distance}$
- Distance: $(|\mathbf{R}_I \mathbf{R}_J| d_0)^2$, $d_0 = \text{target distance}$
- Angle: θ ($\mathbf{R}_I, \mathbf{R}_J, \mathbf{R}_K$)
- Dihedral angle: $\Theta(\mathbf{R}_I, \mathbf{R}_J, \mathbf{R}_K, \mathbf{R}_L)$
- Distance difference: $|\mathbf{R}_I \mathbf{R}_J| |\mathbf{R}_J \mathbf{R}_K|$
- Rigid sub-structure: $\{\mathbf{R}_{I,J,K,...}\}$
- Coordination number: $\mathbf{R}_{I}, \{\mathbf{R}_{J,K,...}\}$

Coordination number

• Coordination number

$$n_I\left(\mathbf{R}^N\right) = \sum_{J\neq I} S\left(|\mathbf{R}_I - \mathbf{R}_J|\right)$$

• Convolution

$$S(r) = \frac{1}{\exp \left[\kappa \left(r - r_c\right)\right] + 1}$$

- κ^{-1} = width of transition region, r_c = coordination radius
- force

$$\mathbf{g}_{I} = -\lambda \frac{\partial n_{I}}{\partial \mathbf{R}_{I}} = \frac{1}{2} \frac{\lambda \kappa}{\cosh\left[\kappa \left(r - r_{c}\right)\right] + 1} \frac{\mathbf{R}_{JI}}{R_{JI}}$$

acts only within the vicinity of r_c

Coordination number



Rare event: Barrier crossing

• It takes a very long time to cross high barriers if the height is much larger than $k_B T$



Rare events

- MD is a real time method, with a time step of the order of 0.1–1 fs
- however, in Nature many effects occur at time scales much longer than the realistic times (*e. g.* in biology seconds)
- due to high energy barriers or inprobable location in phase space (in Arrhenius rate of reaction low prefactor)
- Ways to direct reactions:
 - high temperature
 - constraints
 - bias potentials
 - metadynamics

Free energy differences Thermodynamics integration

- reaction coordinate ζ
- probabilistic interpretation of free energy

$$W\left(\zeta'\right) = -kT \ln P_{\zeta}\left(\zeta'\right)$$
$$P_{\zeta}\left(\zeta'\right) = \left\langle \delta \left[\zeta \left(\mathbf{R}^{N}\right) - \zeta'\right] \right\rangle$$

• Free energy difference

$$W(\zeta_2) - W(\zeta_1) = \int_{\zeta=\zeta_1}^{\zeta_2} \left\langle \frac{\partial \mathcal{H}}{\partial \zeta} \right\rangle_{\zeta'}^{\text{conditional}} d\zeta$$

Blue moon ensemble Reaction rates

Principle:

- Run constrained dynamics (reaction coordinate ζ) from the reactants (ζ_A) to the transition state (ζ^{\ddagger}) \Rightarrow "equilibrium part of rate constant"
- Normal (unconstrained) MD from the constrained MD configuration at the transition state $\zeta^{\ddagger} \Rightarrow$ "dynamic" part of the rate constant": Will land on A or B



• Total rate constant: Product of those two

Blue moon ensemble Example: S_N2 reaction F⁻+CH₃Cl \rightarrow CH₃F + Cl⁻

Mugnai, Cardini & Schettino, JCP 2003

- Left: (Free) energy profile along the reaction paths at 0 and 300 K
- Right: Dipole moment CH_3X and Y^- along the 0 K reaction path





Rare events Complicated energy surfaces



- What happenes if a single reaction coordinate is not enough?
- The low-energy path might not be captured

Rare events



• ... or what if the potential energy landscape is very rugged?

Or: Throwing ropes over rough mountains passes, in the dark

- Transition path sampling, idea:
 - Start with an initial reaction path
 - At a point along the present path start MD with velocities $\dot{\mathbf{R}}_I$ and $-\dot{\mathbf{R}}_I$; if they lead to different end states, the state is probably close to a transition state



 This way new reaction paths can be found, with a lower reaction barrier: Starting from a point "p" we might find another path with a lower barrier



• Reaction rates can be obtained, but with a large amount of statistics (different MD runs)

Transition path sampling Example: 7-atom Lennard-Jones clusters

David Chandler & co-workers

- Move the central (red) atom to the boundary; the central atom in the final state random (blue)
- Difficult, correlated "reaction" mechanism



Transition path sampling Example: Hydrogenated water trimer

• Reaction path found using transition path sampling (left) and the (a) high-energy and (b) low-energy saddle point



Metadynamics

Alessandro Laio & Michele Parrinello, PNAS (2002)

- A method to "drive" chemical reactions using *collective variables*
- Add a small, repulsive potential at the present value of the reaction coordinate
- Free energy surface can be reconstructed after the simulation



• Choose a se of *collective* variables, *e.g.* distances, coordination number, simulation cell parametres, . . .

 $- s_i = s_i \left(\{ \mathbf{R}_I \}_{\text{subset}} \right)$

 $\bullet\,$ Constraint these collective variables at a given point in ${\bf s}\,$

 $- \mathcal{L} = \mathcal{T} - \mathcal{V} + \sum_{i} \lambda_{i} \left| s_{i} \left(\{ \mathbf{R}_{I} \}_{\text{subset}} \right) - s_{i}^{t} \right|$

- Perform "metadynamics" in space of collective coordinates...
 - either in steps: "coarse grained dynamics"
 - continuously: "smooth metadynamics"

Metadynamics

• History-dependent potential

$$V(t,\mathbf{s}) = \int_{t'=0}^{t} \left| \dot{\mathbf{s}}(t') \right| W(t') \exp\left\{ -\frac{\left[\mathbf{s} - \mathbf{s}(t')\right]^2}{2\left(\Delta s^{\perp}\right)^2} \,\delta\left(\frac{\dot{\mathbf{s}}(t')}{\left|\dot{\mathbf{s}}(t')\right|} \cdot \left[\mathbf{s} - \mathbf{s}(t')\right] \right) \right\} \, dt'$$

































- The free energy surface can be reconstructed afterwards!
- $F(\mathbf{s}) = -k_B T \ln P((s)), P((s)) = \frac{1}{Q} \int_{\mathbf{s}'} \exp\left[-E(\mathbf{s})/(k_B T)\right] \delta(\mathbf{s} \mathbf{s}') d\mathbf{s}'$
- Slowly all the local minima are filled and



Metadynamics: Applications so far Rapidly expanding

System (collective variables):

- Solid state phase transitions (cell parametres): graphite to diamond under pressure; silicon; benzene; Li-ABW zeolite (Li [AlSiO₄] · H₂O); melting of ice
- Chemical reactions
 - $S_N 2$ reaction $CI^-+CH_3CI \rightarrow CH_3CI + CI^-$ (distances)
 - carbonylation of epoxide (ethylenoxide) to β -lactone, $[CH_2]_2O$ (EtO) + $CO_2 \rightarrow C_3H_4O_2$ (coordination number)
 - azulene-to-naphthalene $(C_{10}H_8)$ rearrangement (coordination number)
 - isomerisation and dissociation of silicon clusters Si_6H_n , n = 4, 6, 8 (distances)
 - dethreading of cyclophane (distances, coordination number)

Mostly by Michele Parrinello & co-workers

Metadynamics: Applications continue Dethreading of cyclophane

Tetracationic cyclophane (cyclobis-(paraquat-p-phenylene)₈⁴⁺), 1,5-dihydroxynaphthalene, solvated in acetonitrile (distances, coordination number)





Advantages:

- General
- Can cope with high dimensionality
- Predictive, wide exploration of free energy surface (with lower resolution)

Disadvantages:

- Careful choice of the collective variables
- Inaccurate if a "slow" variable is forgotten (can be checked a posteriori)
- Choice of good (optimal) parametres (masses, coupling constants, ...) not straightforward

Molecular dynamics: Summary

- Temperature can be controlled via rescaling (initial) equilibration and thermostats (*e. g.* Nosé-Hoover thermostat chains) for NVT ensemble
- Constraints can be used to pose restrictions on the atoms
- They can be used to direct reactions, however in complicated (potential/free) energy landscapes they might not yield the correct reaction path (in reasonable simulation time, at least)
- Metadynamics looks like a promising method for finding reaction paths and (potential/free) energy surfaces