Molecular dynamics

Molecular dynamics

Why?

- allows realistic simulation of equilibrium and transport properties in Nature
- ensemble averages can be used for statistical mechanics
- time evolution of chemical reactions, phase transitions, . . . can be followed
- search for reaction paths, exploration of phase space

Molecular dynamics

What?

- propagation of Newton's equation of motion (with discrete equations of motion)

$$\mathbf{F}_I = M_I \mathbf{a} = M_I \ddot{\mathbf{R}}_I$$

- alternative derivation from the Lagrange formalism:

$$\mathcal{L}\left(\mathbf{R}^N, \dot{\mathbf{R}}^N\right) = \sum_{I=1}^{N} \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 - U\left(\mathbf{R}^N\right) \ ,$$

  U is the interaction potential between the particles. The Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{R}}_I} = \frac{\partial \mathcal{L}}{\partial \mathbf{R}_I}$$

- most common algorithm: Verlet algorithm (in a few variations)

Verlet algorithm

Velocity Verlet

- discretisation of Newton's equation of motion

$$M_I\ddot{\mathbf{R}}_I = \mathbf{F}_I$$

i) Propagate ionic positions $\mathbf{R}_I(t)$ according to

$$\mathbf{R}_I(t+\Delta t) = \mathbf{R}_I(t) + \Delta t\ \mathbf{v}_I(t) + \frac{(\Delta t)^2}{2M_I}\mathbf{F}_I(t)$$

ii) Evaluate forces $\mathbf{F}_I(t+\Delta t)$ at $\mathbf{R}_I(t+\Delta t)$

iii) Update velocities

$$\mathbf{v}_I(t+\Delta t) = \mathbf{v}_I(t) + \frac{\Delta t}{2M_I}\left[\mathbf{F}_I(t) + \mathbf{F}_I(t+\Delta t)\right]$$

Velocity Verlet

Derivation

- Taylor expansion for ionic positions $\mathbf{R}_I(t)$

$$\begin{aligned}\mathbf{R}_I(t+\Delta t) &= \mathbf{R}_I(t) + \Delta t\ \dot{\mathbf{R}}_I(t) + \frac{(\Delta t)^2}{2}\ \ddot{\mathbf{R}}_I(t) + \ldots \\ &= \mathbf{R}_I(t) + \Delta t\ \mathbf{v}_I(t) + \frac{(\Delta t)^2}{2M_I}\ \mathbf{F}_I(t) + \ldots\end{aligned}$$

- Backward Taylor expansion for ionic positions $\mathbf{R}_I(t)$

$$\mathbf{R}_I(t+\Delta t) = \mathbf{R}_I(t) - \Delta t\ \mathbf{v}_I(t+\Delta t) + \frac{(\Delta t)^2}{2M_I}\mathbf{F}_I(t+\Delta t) + \ldots$$

- Add up:

$$\mathbf{R}_I(t+\Delta t)+\mathbf{R}_I(t) = \mathbf{R}_I(t+\Delta t)+\mathbf{R}_I(t)+\Delta t\,[\mathbf{v}_I(t) - \mathbf{v}_I(t+\Delta t)]+\frac{(\Delta t)^2}{2M_I}\,[\mathbf{F}_I(t) + \mathbf{F}_I(t+\Delta t)]$$

- Yields velocities

$$\mathbf{v}_I(t+\Delta t) = \mathbf{v}_I(t) + \frac{\Delta t}{2M_I}\,[\mathbf{F}_I(t) + \mathbf{F}_I(t+\Delta t)]$$

Velocity Verlet

Advantages

Other algorithms provides can have better short time stability and allow larger time steps, but . . .

- simple and efficient; needs only forces, no higher energy derivatives
- still correct up to and including third order, $(\Delta t)^3$
- explicitly time reversible
- sympletic: conserves volume in phase space
- *superior long time stability (energy conservation) of the Verlet algorithm*

Velocity Verlet

Choice of time step

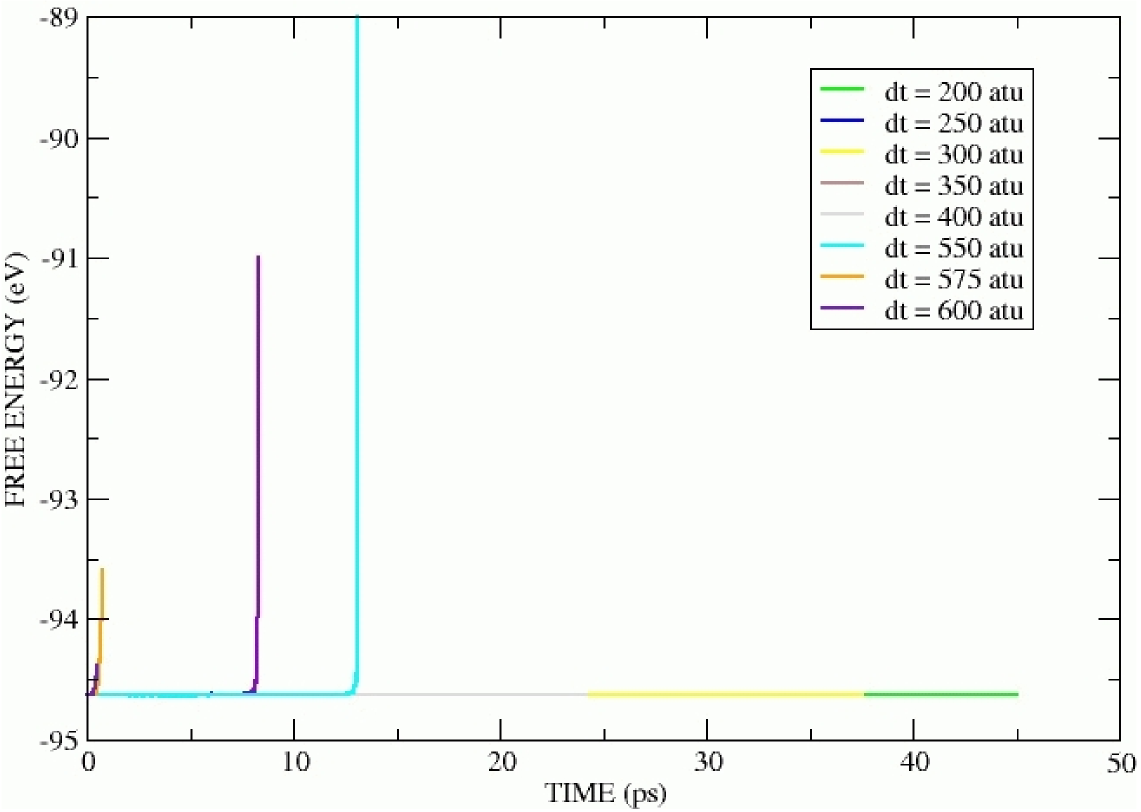
- The time step is in general chosen as large as possible . . .
- "possible" = stable dynamics = energy conserved; or, drift in energy acceptable
- rule of thumb: 6-10 times smaller than the fastest period in the system; otherwise sampling of that mode is impossible
- time step can be changed during simulation(!)

Velocity Verlet: Choice of time step

$AlCl_3$ dimer

Example of a good/bad choice of time step

- Highest vibrational frequency 595 $\text{cm}^{-1} \Rightarrow$ period $T = 56$ fs
- Divergence between $\delta t = 400..500$ atu $= 9.6\text{-}12.0$ fs $\approx 1/5$ T



Equations of motion: Alternative derivation

Propagation methods

- Define phase space vector $\Gamma = (x, p)$ and commutator

$$\{A, H\} = \frac{\partial A}{\partial x}\frac{\partial H}{\partial p} - \frac{\partial A}{\partial p}\frac{\partial H}{\partial x}$$

- Hamilton's equations of motion:

$$\frac{d\Gamma}{dt} = \{\Gamma, H\}$$

- Define $\widehat{\mathcal{L}}$ so that

$$i\widehat{\mathcal{L}}\Gamma = \{\Gamma, H\}$$

- $\dot{\Gamma} = i\widehat{\mathcal{L}}\Gamma \Rightarrow$

$$\Gamma(t) = e^{i\widehat{\mathcal{L}}t}\Gamma(0)$$

- Such formalism has been used by Mark Tuckerman *et al* to derive new integrators

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

$$\left(\sum_{I=1}^{N} \frac{{\mathbf{P}_I}^2}{2M_I} \right)_{\text{NVT}} = \frac{3}{2} N k_B T$$

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

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

- Introduce *instantaneous* temperature

$$\mathcal{T}\left(\mathbf{P}^N\right) = \frac{1}{3 N k_B} \sum_{I=1}^{N} M_I {v_I}^2$$

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

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

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

$$\ddot{\mathbf{R}}_I = \frac{\mathbf{F}_I}{M_I} - \zeta \dot{\mathbf{R}}_I \ ,$$

$$\dot{\zeta} = \frac{1}{Q} \left[\sum_{I=1}^{N} M_I \dot{\mathbf{R}}_I^2 - 3 N k_b T \right]$$

- If ...
  - $\mathcal{T} > k_B T$: Increase of $\dot{\zeta}$, (eventually) larger friction
  - $\mathcal{T} < k_B T$: 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 T
  - couples to the frequencies of the system; usually selected as a characteristic frequency ω_n in the spectrum
- new constant of motion:

$$E^{\text{NVT}}_{\text{conserved}} = E_{\text{k,I}} \left(\{ \dot{\mathbf{R}}_I \} \right) + E_{\text{pot}} \left(\{ \mathbf{R} \} \right) + \frac{Q}{2} \zeta^2 + 3 N k_B T \int_t \zeta \, dt$$

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)

$$
\begin{aligned}
M_I \ddot{\mathbf{R}}_I &= \mathbf{F}_I - M_I \dot{\xi} \dot{\mathbf{R}}_I \\
Q_1^n \ddot{\xi}_1 &= \left[\sum_{I=1}^{N} M_I \dot{\mathbf{R}}_I^2 - 3 N k_b T \right] - Q_1^n \dot{\xi}_1 \dot{\xi}_2 \\
Q_k^n \ddot{\zeta}_k &= \left[Q_{k-1}^n \dot{\xi}_{k-1}^2 - k_b T \right] - Q_k^n \dot{\xi}_k \dot{\xi}_{k+1} \left(1 - \delta_{kK} \right) \ , \ k = 2, \ldots, K \\
Q_1^n &= \frac{3 N k_B T}{\omega_n^2} \ , \ Q_k^n = \frac{k_B T}{\omega_n^2}
\end{aligned}
$$

- new energy, but is still conserved:

$$
E_{\text{conserved}}^{\text{NVT}} = E_{\text{k,I}} \left(\left\{ \dot{\mathbf{R}}_I \right\} \right) + E_{\text{pot}} \left(\{ v R \} \right) + \sum_{k=1}^{N} \frac{Q_k^n}{2} \dot{\xi}_k^2 + 3 N k_B T \xi_1 + \sum_{k=2}^{N} k_B T \xi_k
$$

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
- Constant number of particles

Tricks

- Simulated annealing
- Multiple time scales / RESPA
- Periodic boundary conditions
- Ewald summation
- Thermodynamic integration
- Cell lists *etc*

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.* —CH_3 group)

Example: a fixed inter-atomic distance:

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

Constraints

Lagrangean formulation

- Lagrangean (multipliers λ):

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

- Equation of motion:

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

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

Coordination number

- Coordination number

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

- Convolution

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

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

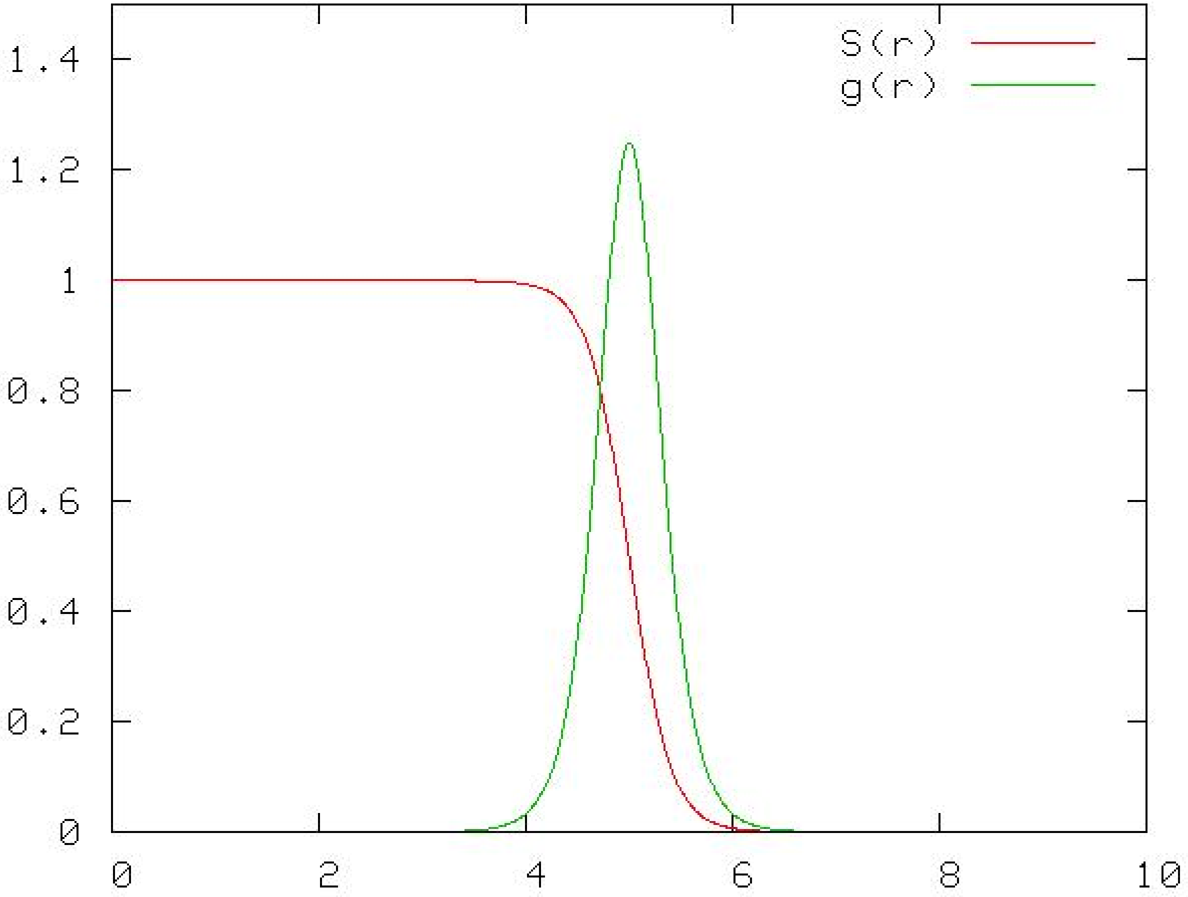
- force

$$\mathbf{g}_I = -\lambda\frac{\partial n_I}{\partial \mathbf{R}_I} = \frac{1}{2}\,\frac{\lambda\kappa}{\cosh\left[\kappa\left(r - r_c\right)\right] + 1}\,\frac{\mathbf{R}_{JI}}{R_{JI}}$$

  acts only within the vicinity of r_c

Coordination number

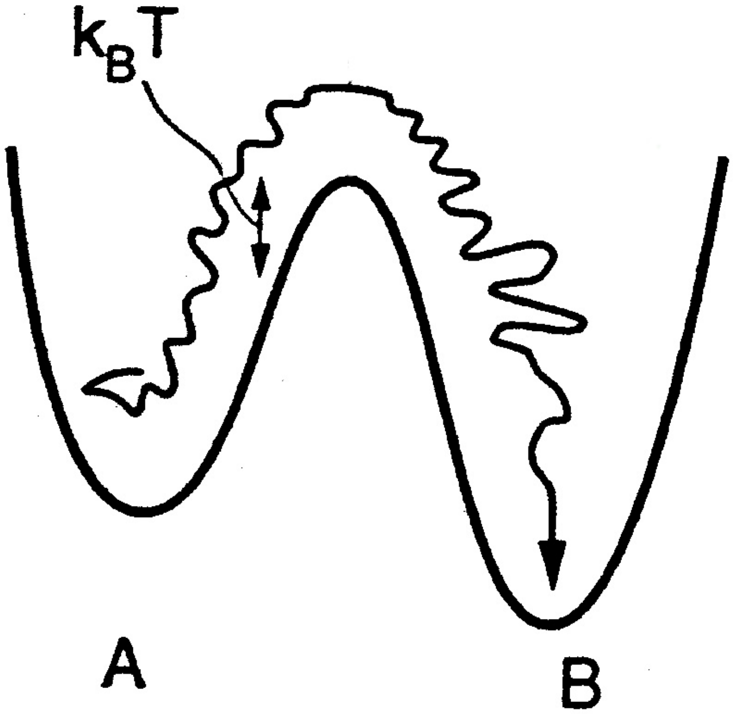
$$S\left(r\right)=\frac{1}{\exp\left[\kappa\left(r-r_c\right)\right]+1} \qquad \mathbf{g}_I=\frac{1}{2}\,\frac{\lambda\kappa}{\cosh\left[\kappa\left(r-r_c\right)\right]+1}\,\frac{\mathbf{R}_{JI}}{R_{JI}}$$



Rare event: Barrier crossing

"To B or not to B"

- 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

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

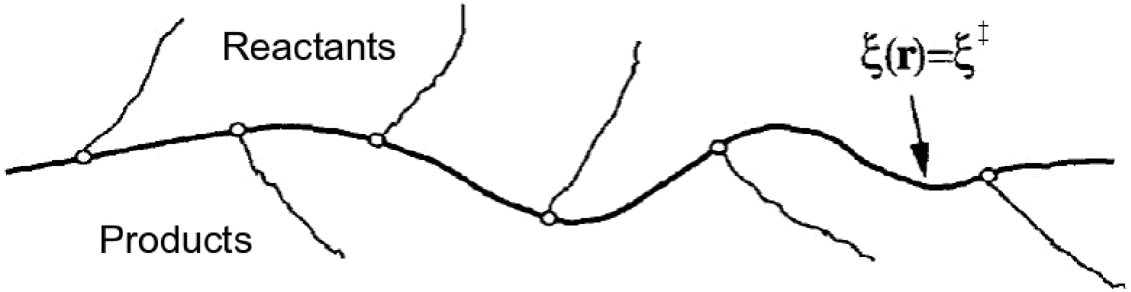
Blue moon ensemble

Reaction rates

Principle:

- Run constrained dynamics (reaction coordinate ζ) from the reactants (ζ_A) to the transition state ($\zeta^\ddagger$) $\Rightarrow$ "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



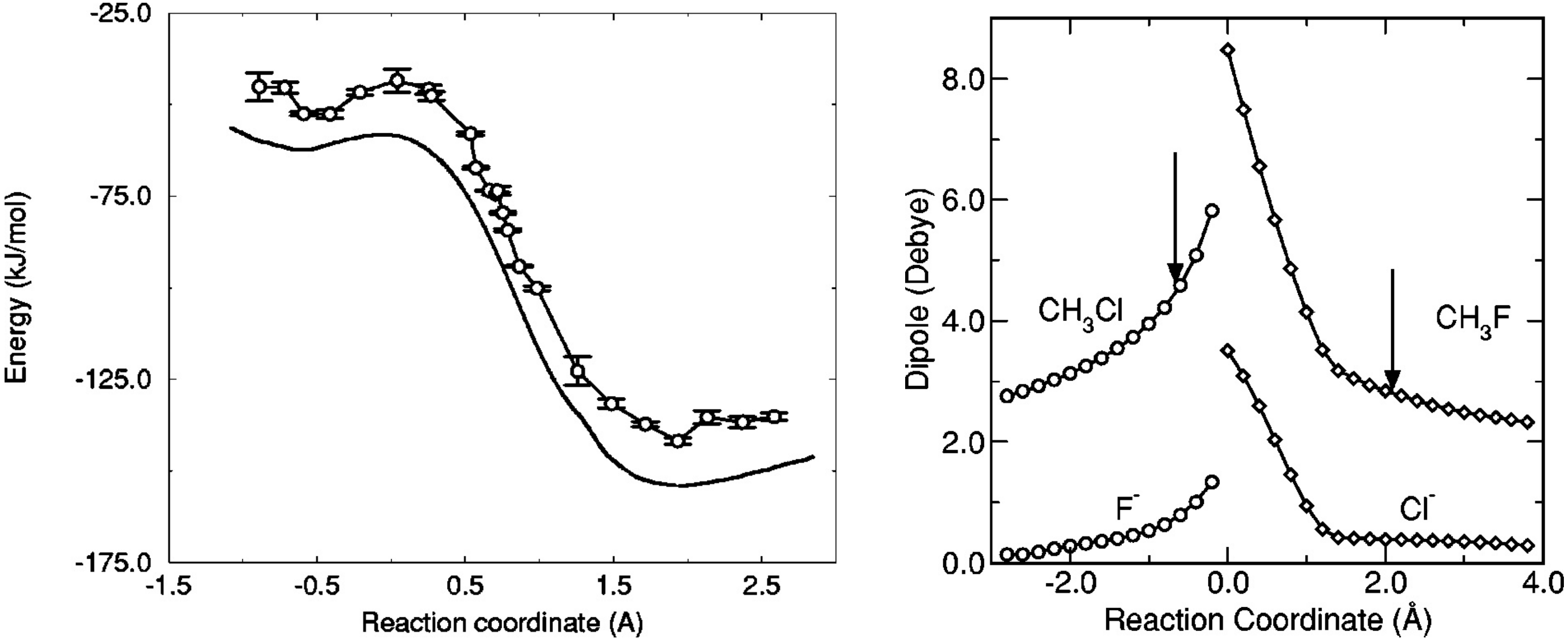
- Total rate constant: Product of those two

Blue moon ensemble

Example: S_N2 reaction $F^{-}+CH_3Cl \rightarrow CH_3F + Cl^{-}$

Mugnai, Cardini & Schettino, *JCP* 2003

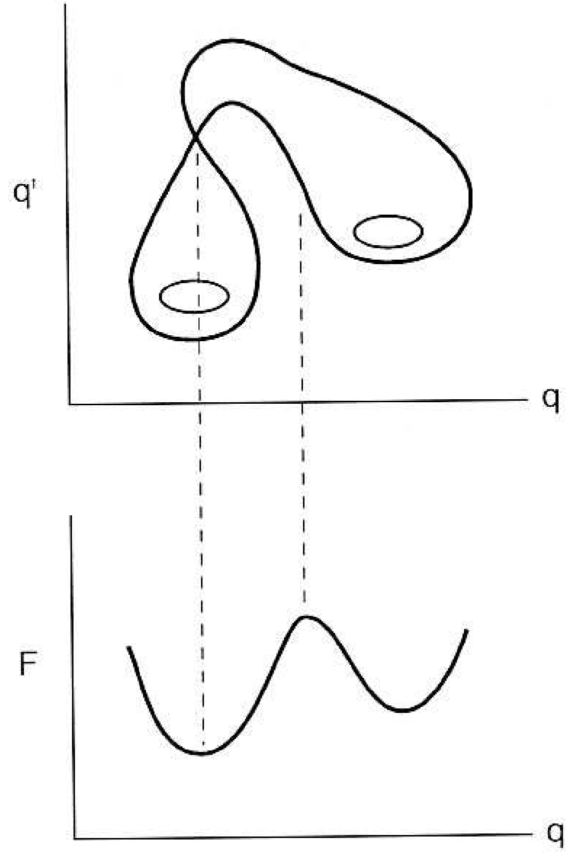
- Left: (Free) energy profile along the reaction paths at 0 and 300 K
- Right: Dipole moment CH_3X and Y^{-} along the 0 K reaction path



But:

Rare events

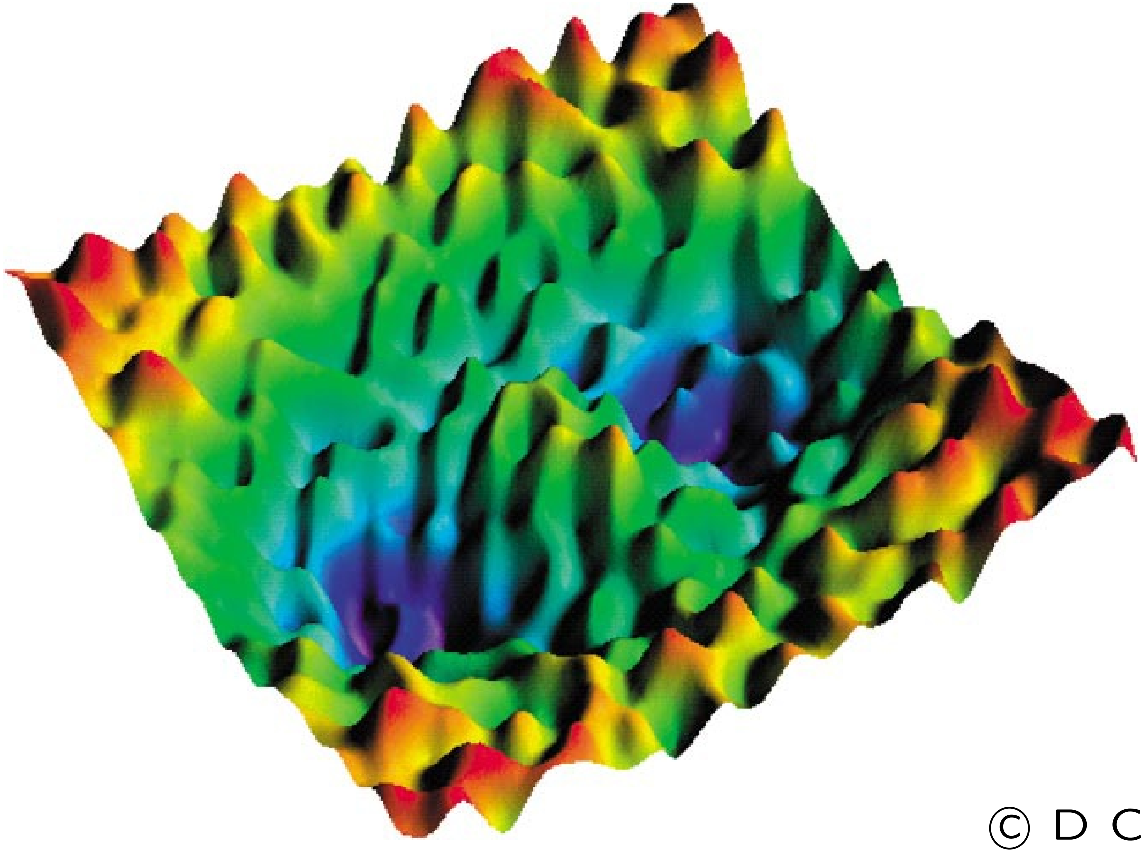
Complicated energy surfaces



© M Parrinello

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

Rare events



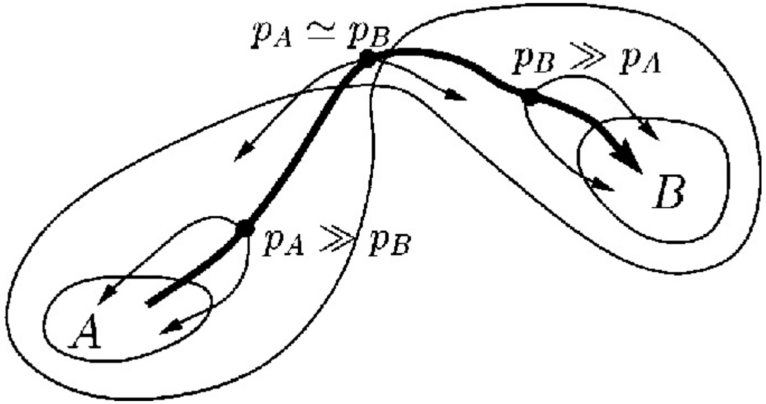
© D Chandler

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

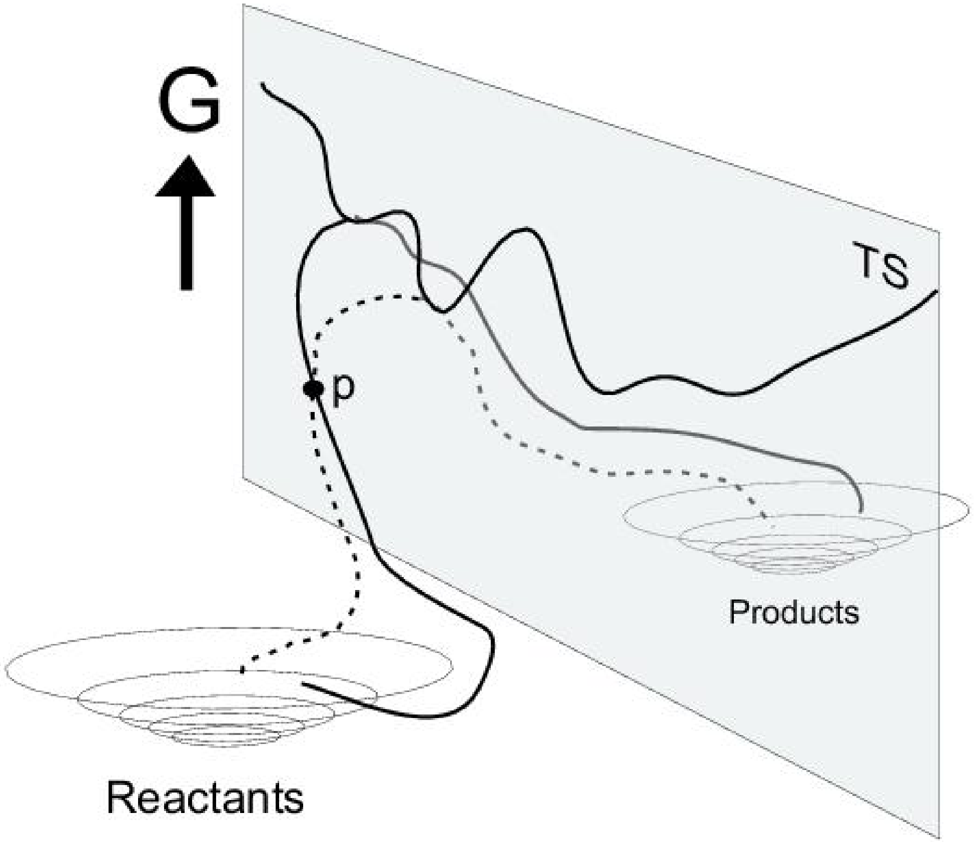
Transition path sampling

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)

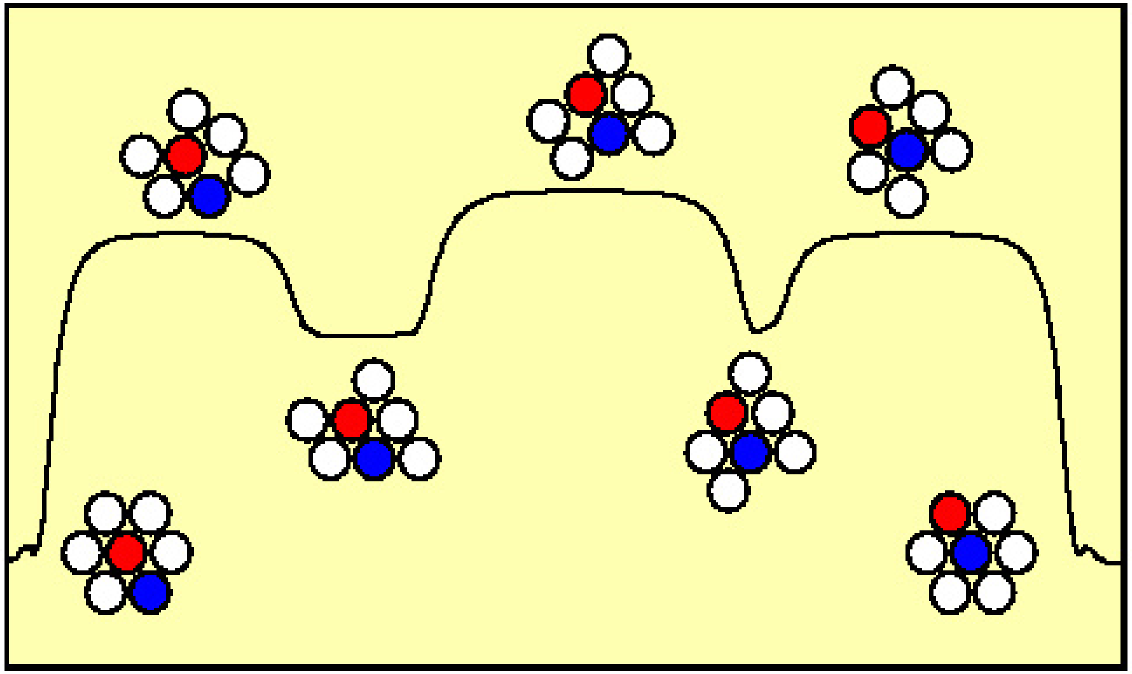
Transition path sampling

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)

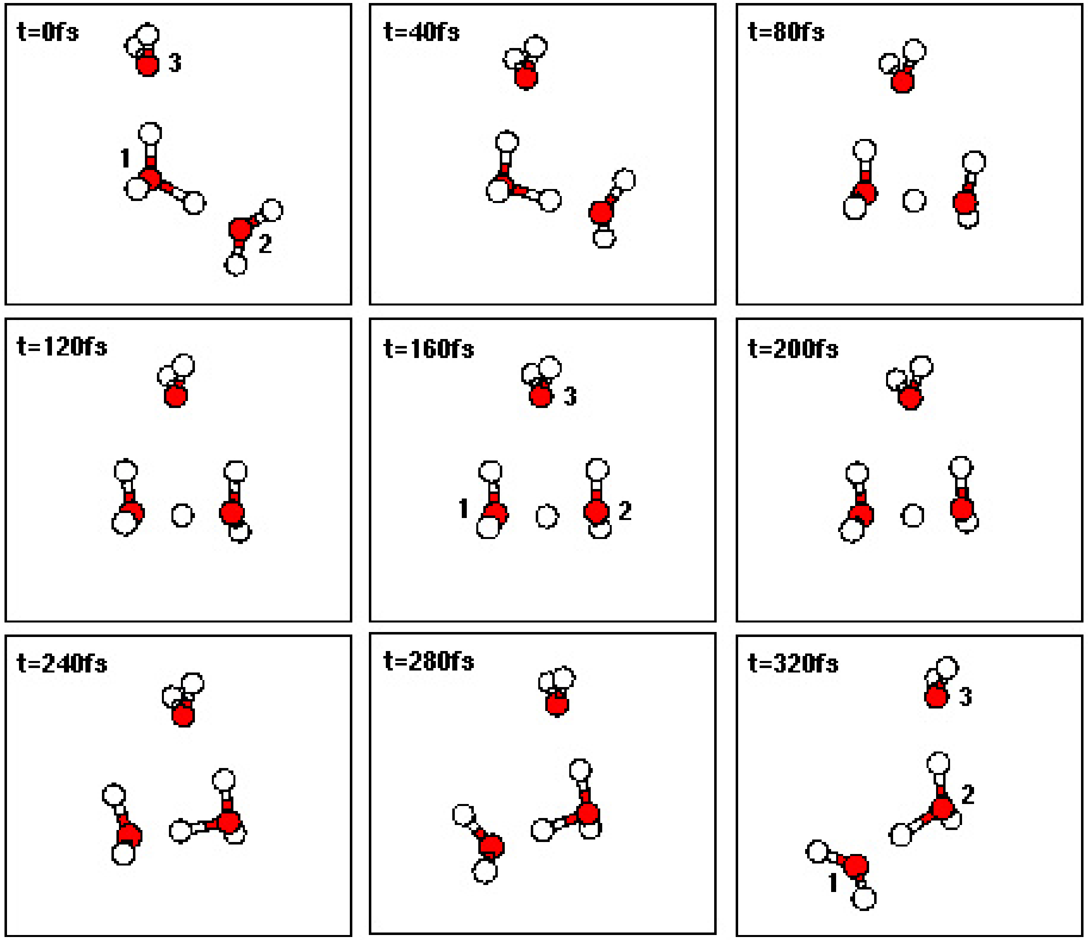
- Difficult, correlated "reaction" mechanism

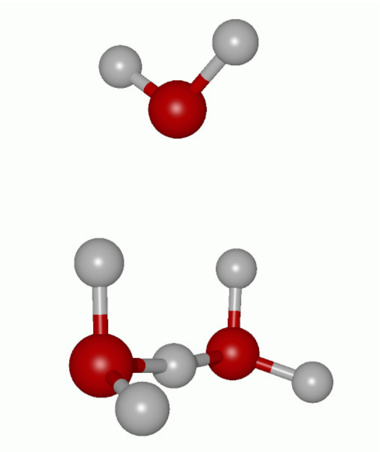


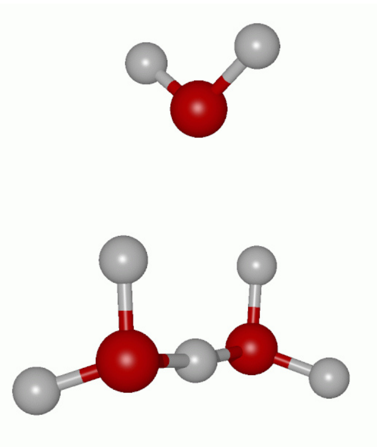
Transition path sampling

Example: Hydrogenated water trimer

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




(a)


(b)

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

Metadynamics

Algorithm

- Choose a se of *collective* variables, *e. g.* distances, coordination number, simulation cell parametres, . . .
  - $s_i = s_i\left(\{\mathbf{R}_I\}_{\text{subset}}\right)$
- Constraint these collective variables at a given point in $\mathbf{s}$
  - $\mathcal{L} = \mathcal{T} - \mathcal{V} + \sum_i \lambda_i \left| s_i\left(\{\mathbf{R}_I\}_{\text{subset}}\right) - s_i^t \right|$
- Perform "metadynamics" in space of collective coordinates. . .
  - either in steps: "coarse grained dynamics"
  - continuously: "smooth metadynamics"

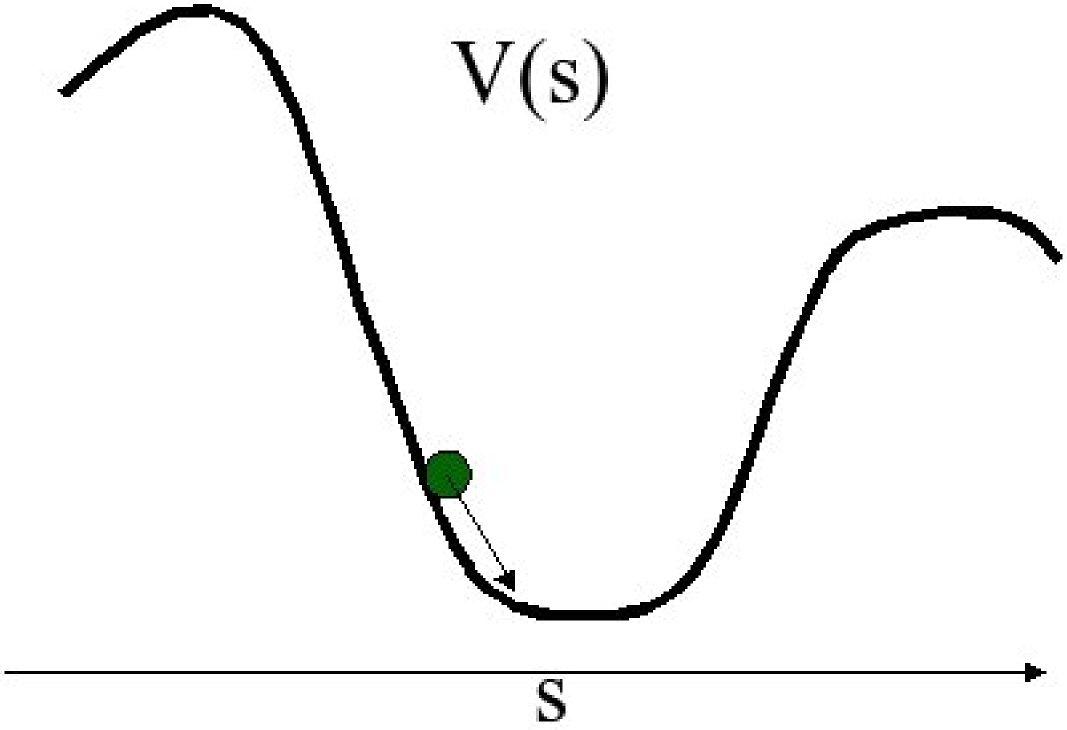
Metadynamics

- History-dependent potential

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

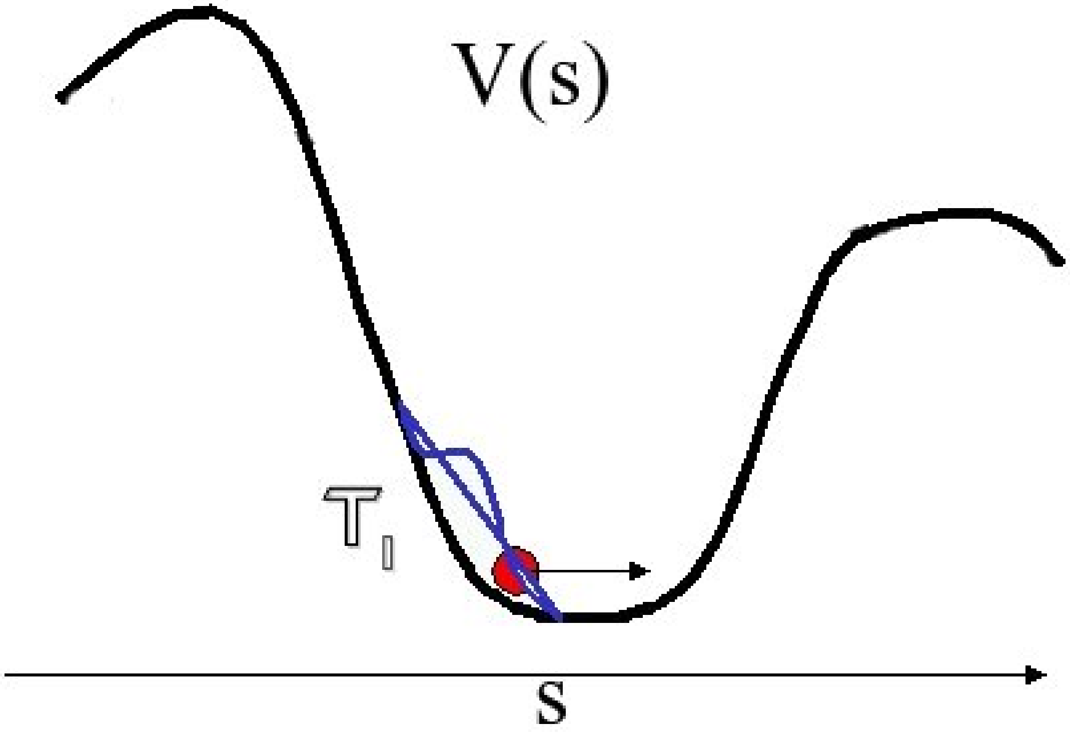
Metadynamics

Evolution of $V(t, s)$



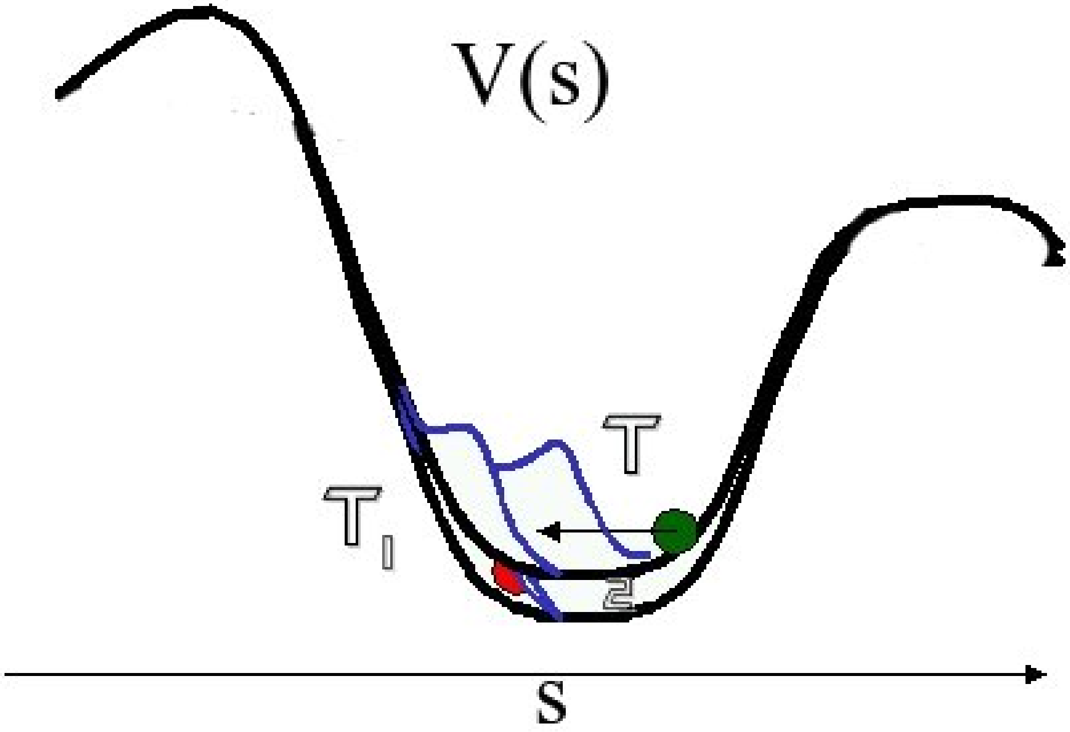
Metadynamics

Evolution of $V(t, s)$



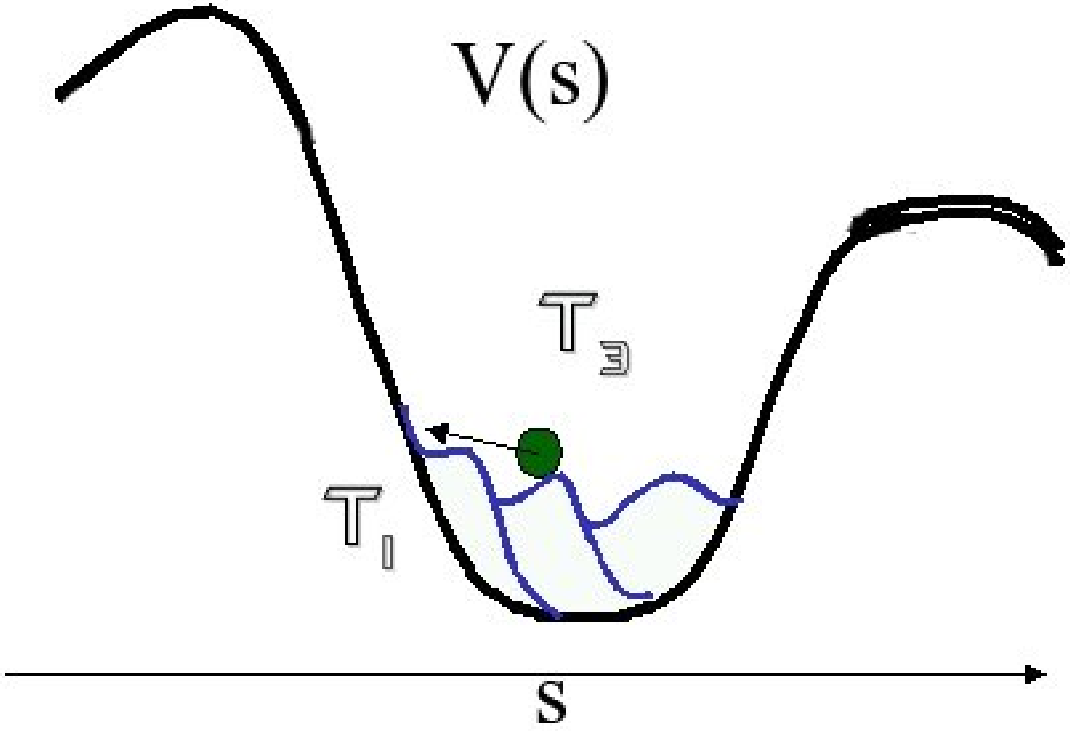
Metadynamics

Evolution of $V(t, s)$



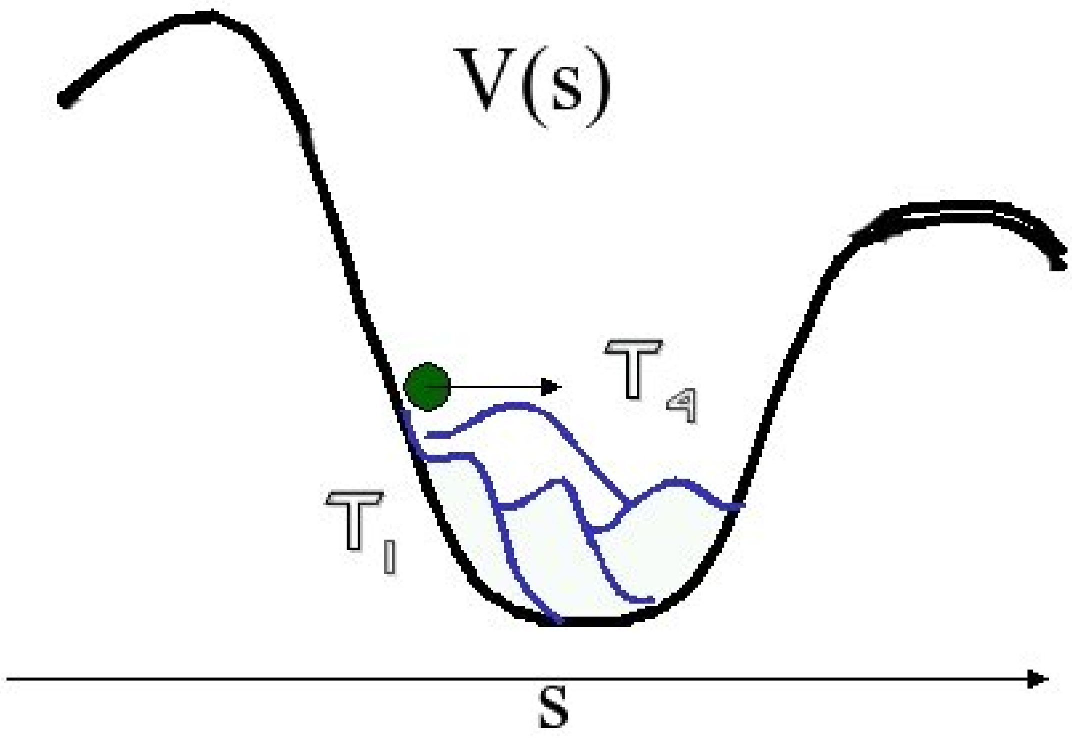
Metadynamics

Evolution of $V(t, s)$



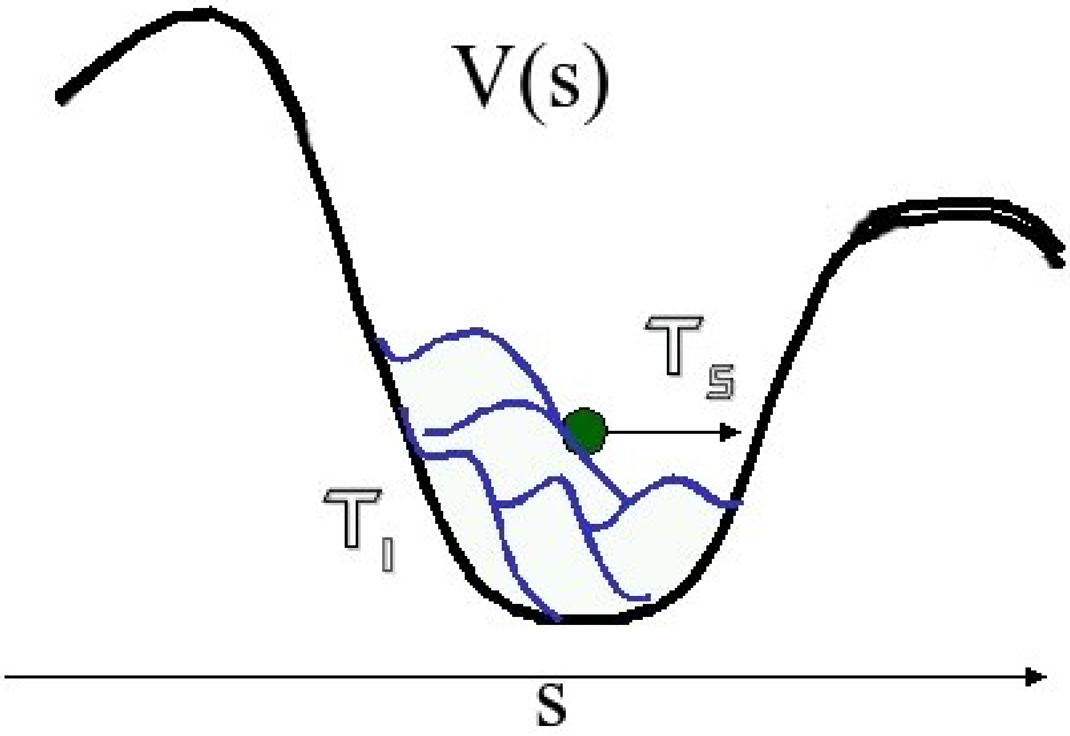
Metadynamics

Evolution of $V(t, s)$



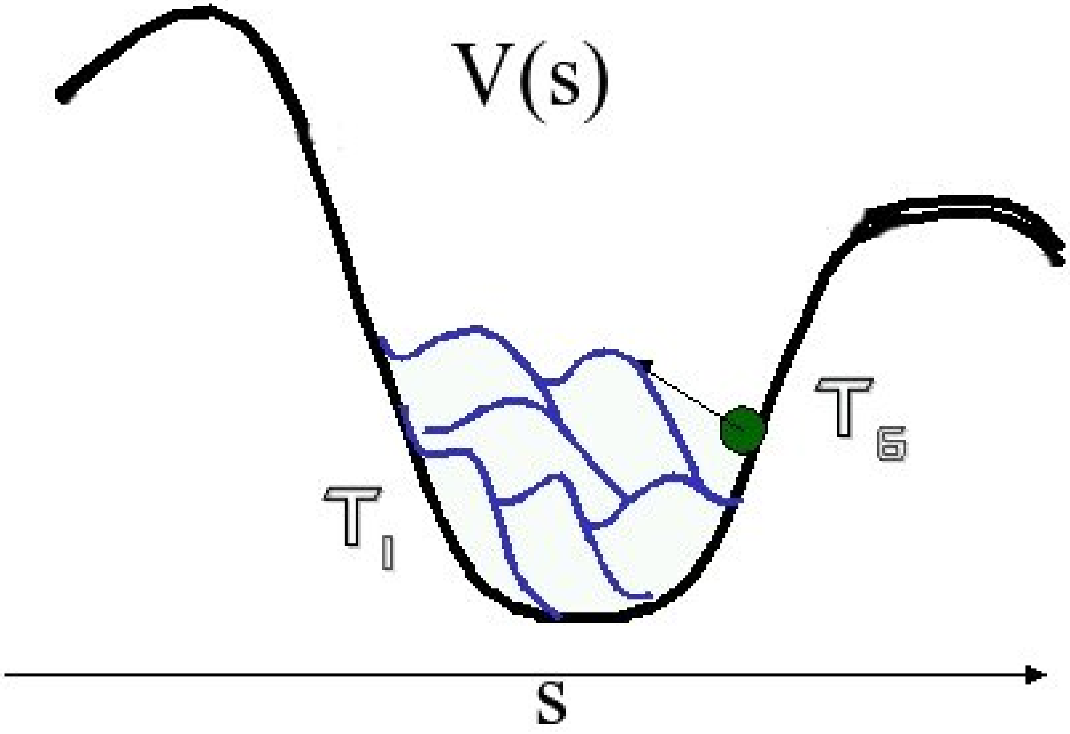
Metadynamics

Evolution of $V(t, s)$



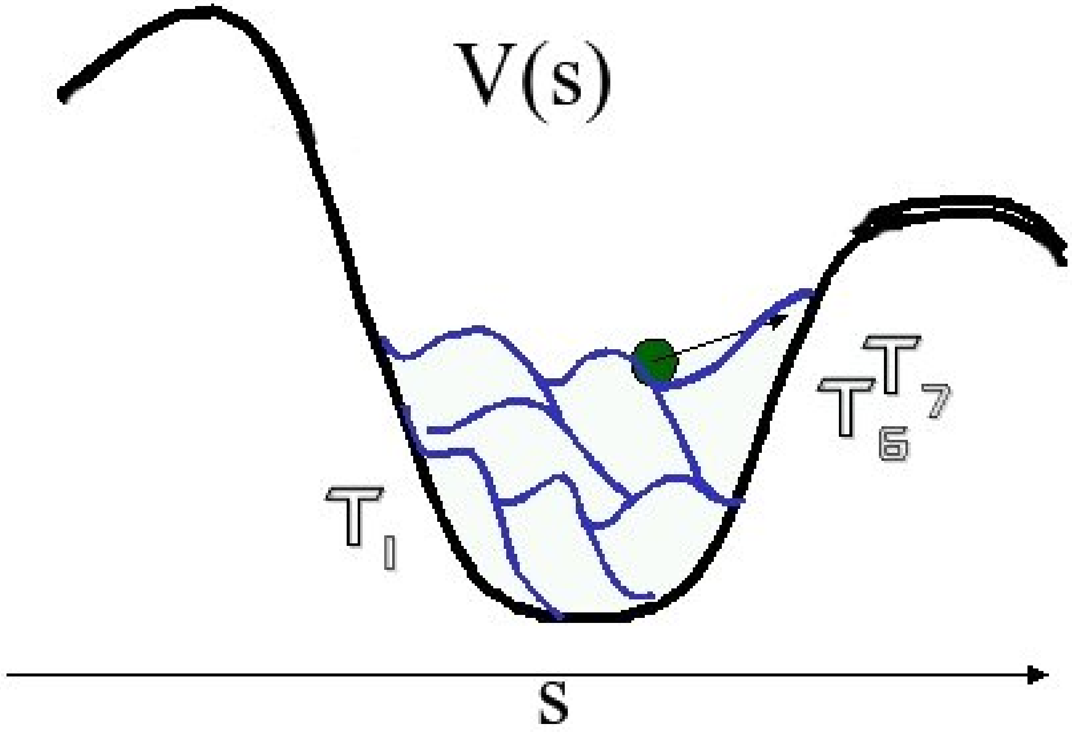
Metadynamics

Evolution of $V(t, s)$



Metadynamics

Evolution of $V(t, s)$

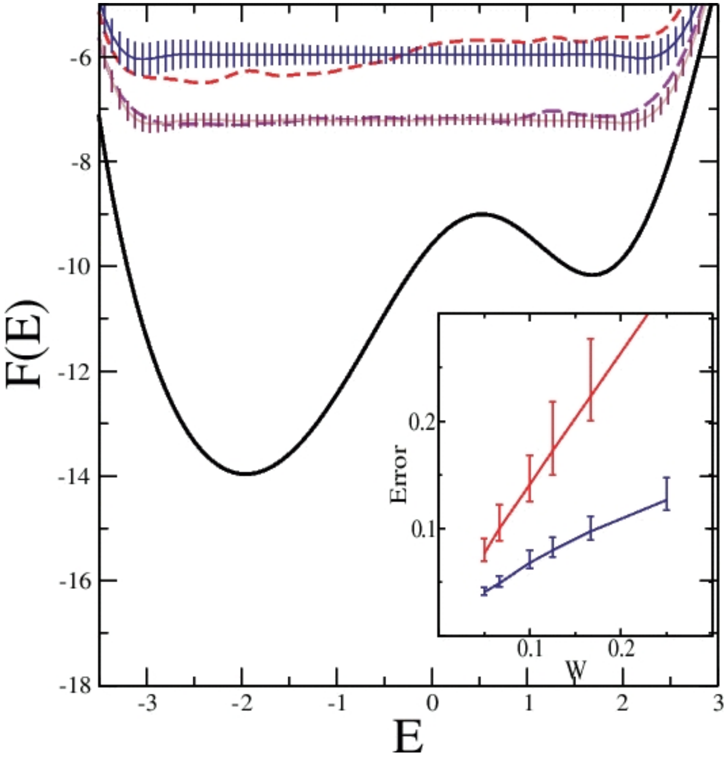


Metadynamics

Free energy surface

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

$$\lim_{t\to\infty} V(t,\mathrm{s}) + F(\mathrm{s}) = constant$$



Metadynamics: Applications so far

Rapidly expanding

System (collective variables):

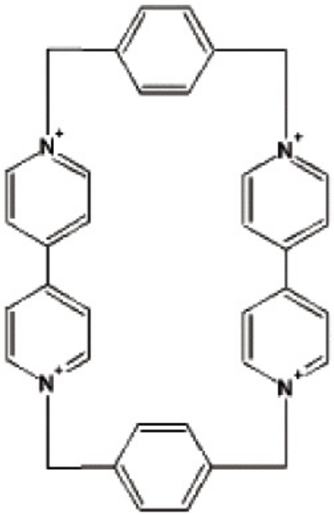
- Solid state phase transitions (cell parametres): graphite to diamond under pressure; silicon; benzene; Li-ABW zeolite ($\mathrm{Li}\,[\mathrm{AlSiO_4}]\cdot\mathrm{H_2O}$); melting of ice
- Chemical reactions
  - S_N2 reaction $Cl^-+CH_3Cl \rightarrow CH_3Cl + Cl^-$ (distances)
  - carbonylation of epoxide (ethylenoxide) to β-lactone, $[CH_2]_2\,O$ (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)

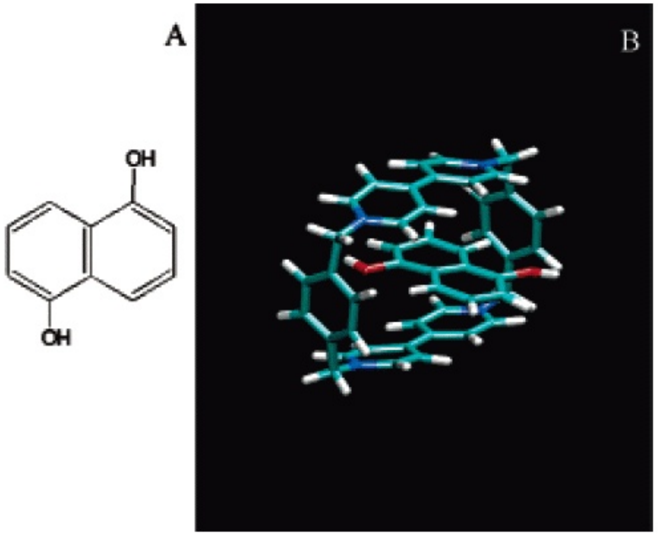
All by Michele Parrinello & co-workers

Metadynamics: Applications continue

Dethreading of cyclophane

Tetracationic cyclophane (cyclobis-(paraquat-p-phenylene)$_8{}^{4+}$), 1,5-dihydroxynaphthalene, solvated in acetonitrile (distances, coordination number)





All by Michele Parrinello & co-workers

Metadynamics

Summary

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

Potentials

(Topic of next talks)

- Empirical classical potentials
  - pair potentials, three-body potentials
  - polarisable force fields
  - effective medium theory, embedded atom method
- Empirical quantum mechanical potentials
  - tight binding Hamiltonian
  - semi-empirical quantum chemistry methods
- *Ab initio* potentials
  - quantum chemistry, methods based on wave function
  - *density functional theory*

Molecular dynamics: Summary

- Molecular dynamics can be used to perform real-time dynamics in atomistic systems

- Verlet algorithm yields stable dynamics (in `CPMD` implemented algorithm velocity Verlet)

- Maximum time step $\Delta t \approx 1$ fs (highest ionic frequency $2000 - 3000$ cm^{-1})

- 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