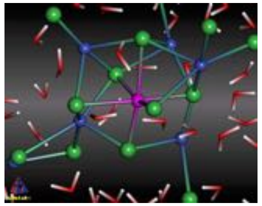
Rare Event Simulations

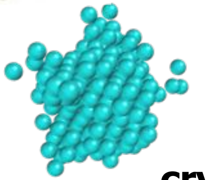
Transition state theory 16.1-16.2
Bennett-Chandler Approach 16.2
Transition path sampling16.4

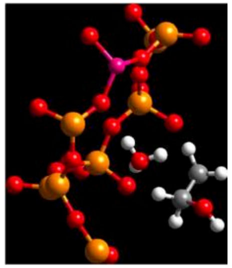
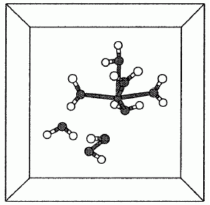
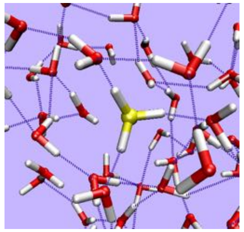
Outline

- Part 1
    - Rare event and reaction kinetics
    - Linear Response theory
    - Transition state theory
    - Free energy methods
    - Bennet Chandler approach
    - Example zeolites

- Part 2
    - Two ended methods
    - Transition path sampling
    - Rate constants
    - Reaction coordinate analysis
    - Application to crystallization
    - Path metadynamics

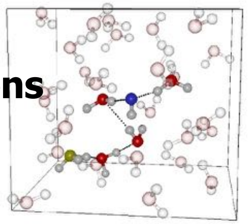


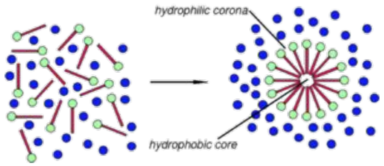


crystallisation

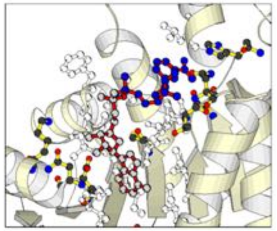
catalysis solution reactions

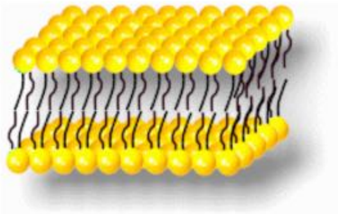




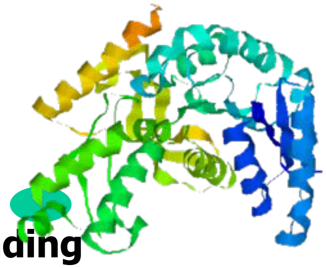
Rare Events

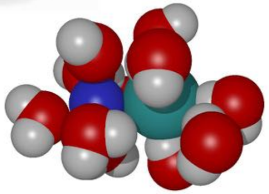
enzyme reactions

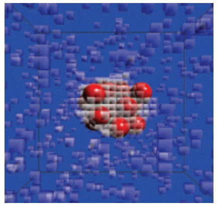




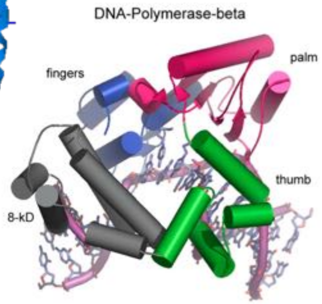
complex fluids





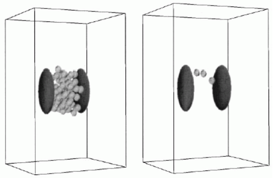
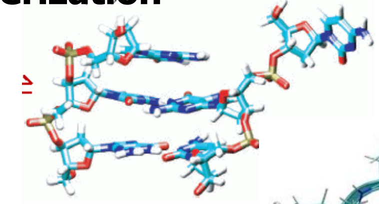
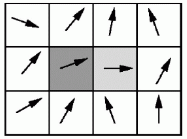
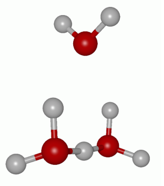
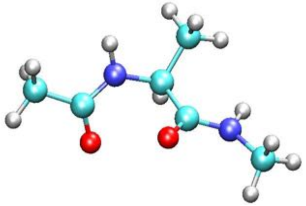
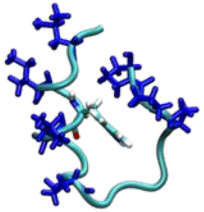
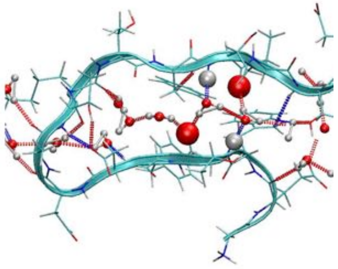


solvent effects



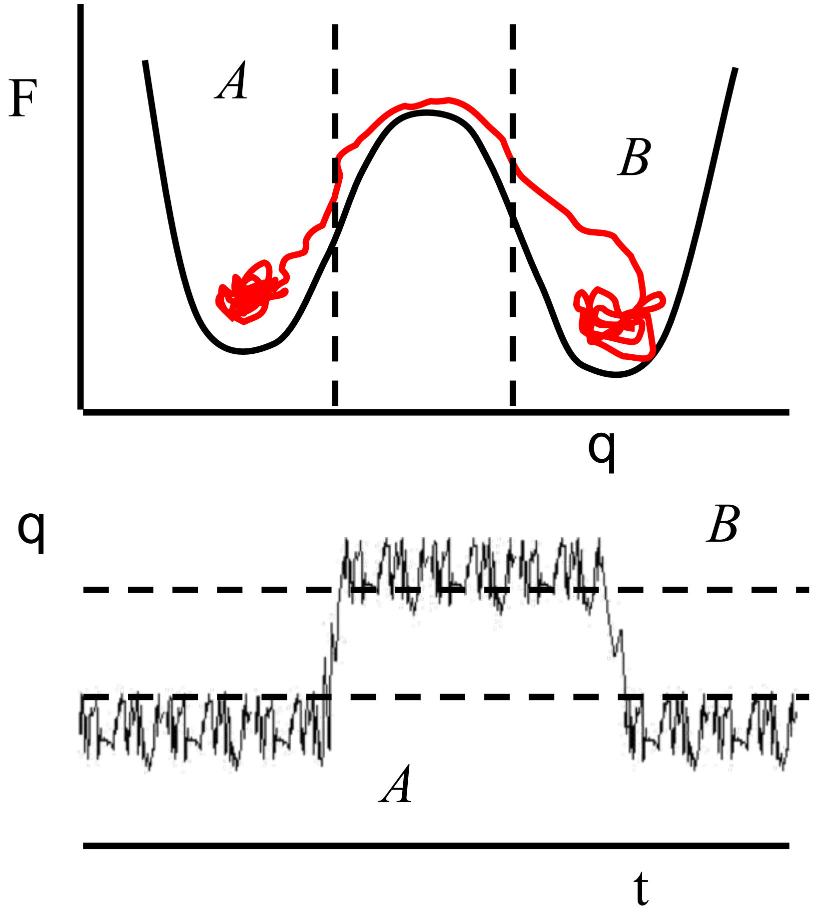
folding & binding

isomerization

Rare events

Interesting transitions in complex systems
- solution chemistry
- protein folding
- enzymatic reactions
- complex surface reactions
- diffusion in porous media
- nucleation



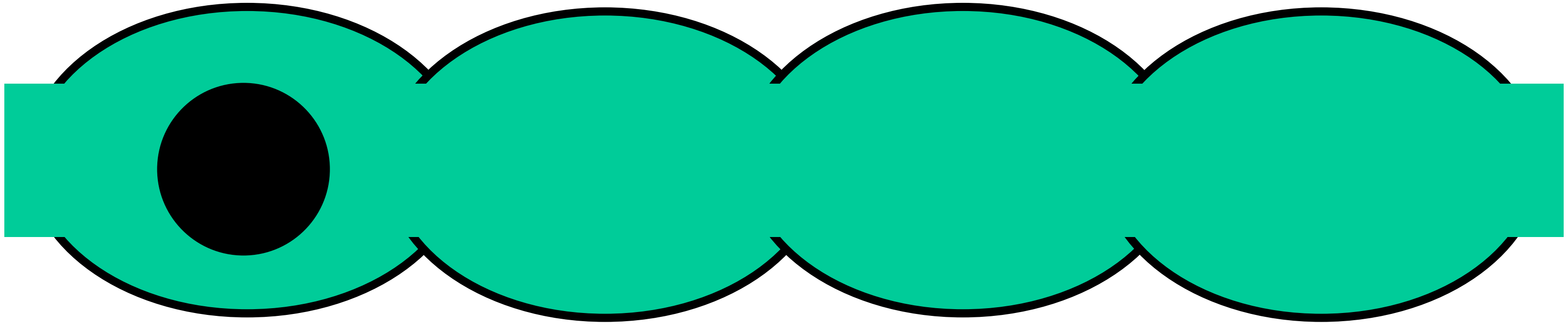
These reactions happen on a long time scale compared to the molecular timescale

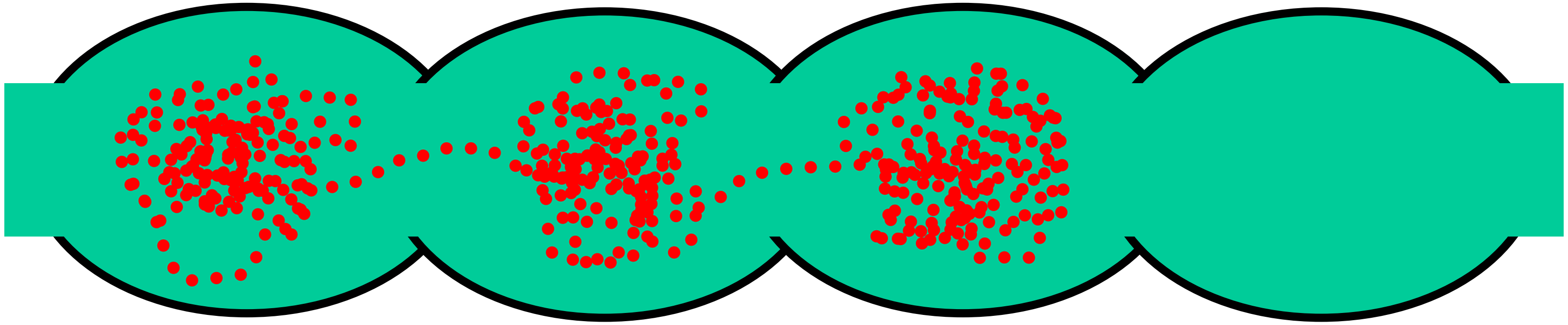
↓

dominated by collective, rare events
Straightforward MD very inefficient

$$\tau_{mol} \ll \tau_{stable}$$

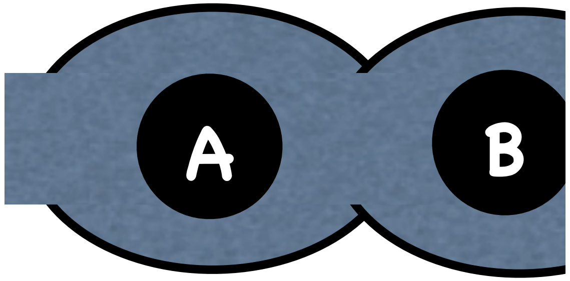
Example: Diffusion in porous material





Phenomenological reaction kinetics

A rare event can be seen as a chemical reaction between reactant A and product B



$$A \leftrightarrow B$$

The change in population c(t) is (0<c<1)

$$\frac{\mathrm{d}c_A(t)}{\mathrm{dt}} = -k_{A\to B}c_A(t) + k_{B\to A}c_B(t) \qquad \frac{\mathrm{d}c_B(t)}{\mathrm{dt}} = +k_{A\to B}c_A(t) - k_{B\to A}c_B(t)$$

Total number change in population

$$\frac{\mathrm{d}\left[c_A(t) + c_B(t)\right]}{\mathrm{dt}} = 0$$

Equilibrium:

$$\dot{c}_A(t) = \dot{c}_B(t) = 0$$

This gives a relation between equilibrium population and reaction rates

$$\frac{\langle c_A \rangle}{\langle c_B \rangle} = \frac{k_{B\to A}}{k_{A\to B}}$$

Relaxation time

Let us make a perturbation of the equilibrium populations, e.g by applying an external field.

When releasing the field, the system will relax to the original equilibrium

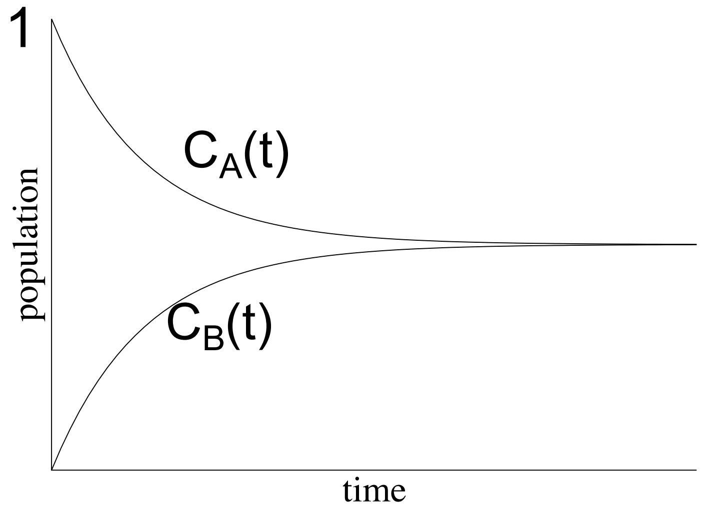
For state A $c_A(t) = \langle c_A \rangle + \Delta c_A(t)$ For state B: $c_B(t) = \langle c_B \rangle - \Delta c_A(t)$

We can rewrite the kinetics in terms of the perturbation Δc:

$$\frac{\mathrm{d}\Delta c_A(t)}{\mathrm{dt}} = -k_{A\to B}\Delta c_A(t) - k_{B\to A}\Delta c_A(t)$$

$$\Delta c_A(t) = \Delta c_A(0)\exp\left[-\left(k_{A\to B} + k_{B\to A}\right)t\right]$$

$$= \Delta c_A(0)\exp\left[-t/\tau\right]$$



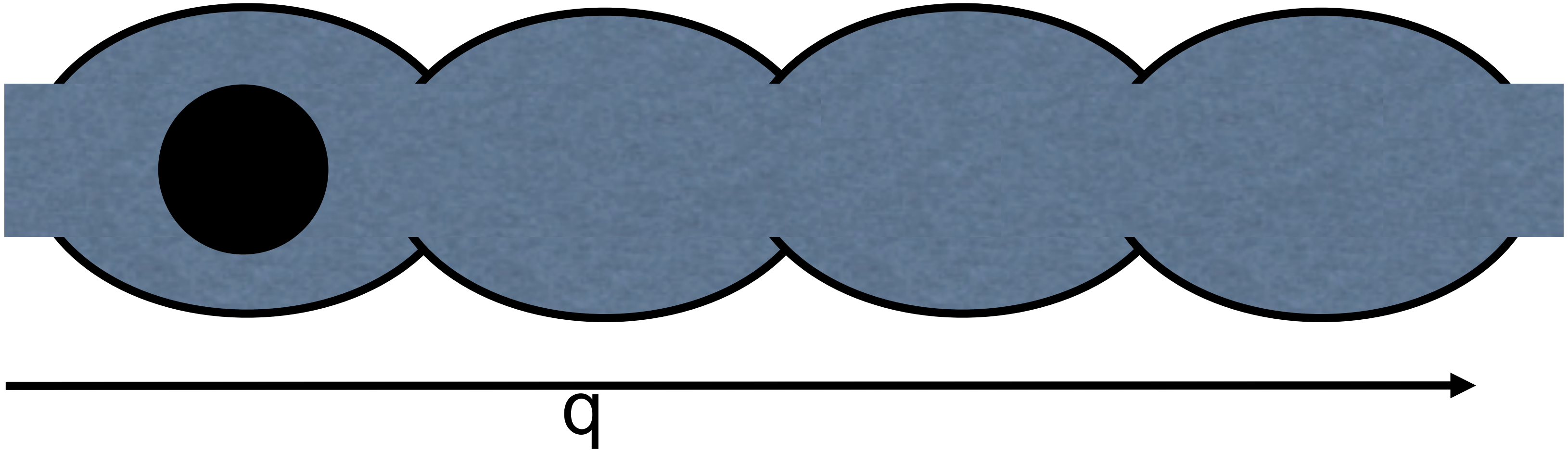
With relaxation time

$$\tau = \left(k_{A\to B} + k_{B\to A}\right)^{-1} = k_{A\to B}^{-1}\left(1 + \langle c_A \rangle / \langle c_B \rangle\right)^{-1} = \frac{\langle c_B \rangle}{k_{A\to B}}$$

$$c_A(t) + c_B(t) = 1$$

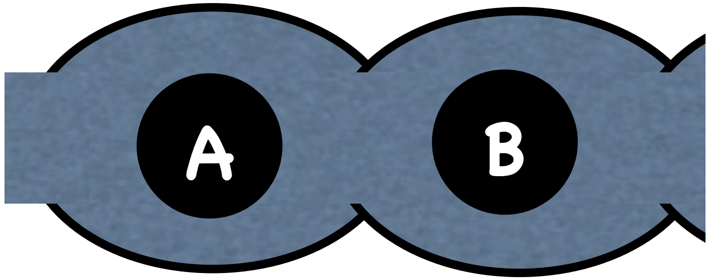
Microscopic theory

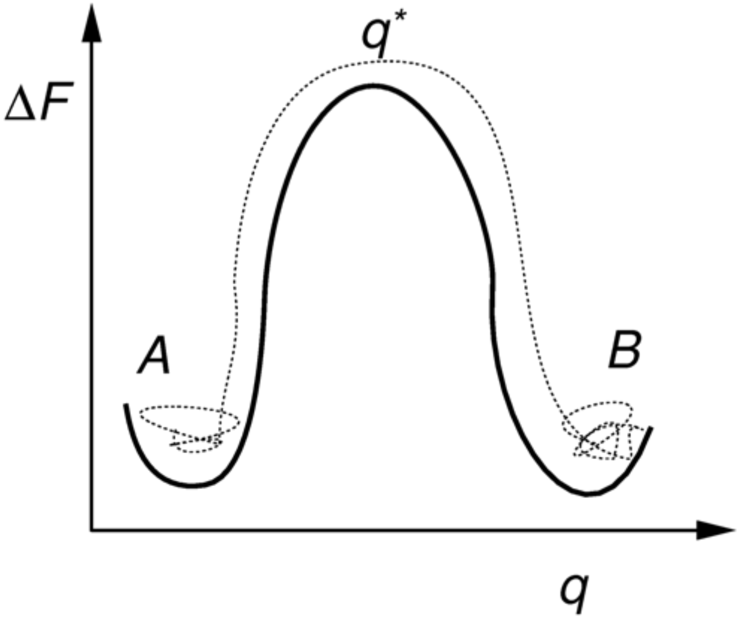
Microscopic description of the progress of a reaction



Reaction coordinate: in this case the z-coordinate of the particle

We need to write the kinetics of the reaction in terms of this microscopic reaction coordinate q





Reaction coordinate

Reactant A: $q < q^*$

Product B: $q > q^*$

Transition state: $q = q^*$

Heaviside θ-function

$$\theta(q-q^*) = \begin{cases} 0 & q-q^* < 0 \\ 1 & q-q^* > 0 \end{cases}$$

Let us introduce the function g_A:

$$g_A(q-q^*) = 1-\theta(q-q^*) = \theta(q^*-q)$$

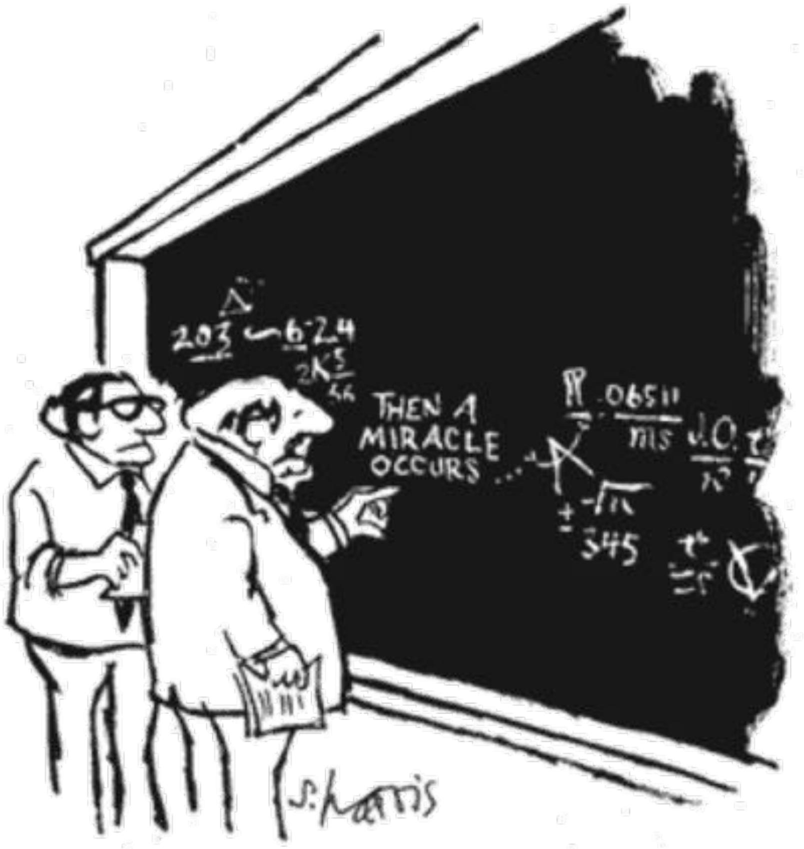
With this function we write for the **probability** $c_A(t)$ the system is in state A:

$$\langle c_A(t)\rangle = \langle g_A(t)\rangle$$

Microscopic theory

Is going to give us the
macroscopic relaxation in terms of a
microscopic time correlation function

$$\exp\left[-t/\tau\right]=\frac{\left\langle \Delta g_A(0)\,\Delta g_A(t)\right\rangle}{\left\langle c_A\right\rangle\left\langle c_B\right\rangle}$$



"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO."

This needs linear response theory

Perturbed Hamiltonian

Let us consider the effect of a static perturbation:

$$H = H_0 - \varepsilon g_A (q - q^*)$$

This external potential increases the concentration of A

For the equilibrium concentration as a function of ε:

$$\Delta c_A = \langle c_A \rangle_\varepsilon - \langle c_A \rangle_0 \quad = \langle g_A \rangle_\varepsilon - \langle g_A \rangle_0$$

We need to compute the ensemble average in the form of :

$$\langle A \rangle_0 = \frac{\int \mathrm{d}\Gamma A \exp[-\beta H_0]}{\int \mathrm{d}\Gamma \exp[-\beta H_0]}$$

Linear Response theory (static)

The original Hamiltonian (H_0) is perturbed by εD:

$$H = H_0 - \varepsilon D$$

This gives as change in the expectation value of A:

$$\langle \Delta A \rangle = \langle A \rangle - \langle A \rangle_0$$

with

$$\langle A \rangle = \frac{\int \mathrm{d}\Gamma A \exp\left[-\beta\left(H_0 - \varepsilon D\right)\right]}{\int \mathrm{d}\Gamma \exp\left[-\beta\left(H_0 - \varepsilon D\right)\right]} \qquad \langle A \rangle_0 = \frac{\int \mathrm{d}\Gamma A \exp\left[-\beta H_0\right]}{\int \mathrm{d}\Gamma \exp\left[-\beta H_0\right]}$$

If the perturbation is small we can write $\langle A \rangle = \langle A \rangle_0 + \frac{\partial \langle A \rangle_0}{\partial \varepsilon} \varepsilon$

For such a small perturbation $$\langle \Delta A \rangle = \frac{\partial \langle A \rangle_0}{\partial \varepsilon} \varepsilon = \left\langle \frac{\partial A}{\partial \varepsilon} \right\rangle_0 \varepsilon$$

with $$\left\langle \frac{\partial A}{\partial \varepsilon} \right\rangle = \frac{\partial \langle A \rangle}{\partial \varepsilon} = \frac{\int \mathrm{d}\Gamma \beta A D \exp\left[-\beta\left(H_0 - \varepsilon D\right)\right] \int \mathrm{d}\Gamma \exp\left[-\beta\left(H_0 - \varepsilon D\right)\right]}{\left\{\int \mathrm{d}\Gamma \exp\left[-\beta\left(H_0 - \varepsilon D\right)\right]\right\}^2}$$
$$- \frac{\int \mathrm{d}\Gamma A \exp\left[-\beta\left(H_0 - \varepsilon D\right)\right] \int \mathrm{d}\Gamma \beta D \exp\left[-\beta\left(H_0 - \varepsilon D\right)\right]}{\left\{\int \mathrm{d}\Gamma \exp\left[-\beta\left(H_0 - \varepsilon D\right)\right]\right\}^2}$$

Evaluated for ε= 0

$$\left\langle \frac{\partial A}{\partial \varepsilon} \right\rangle_0 = \frac{\int \mathrm{d}\Gamma \beta A D \exp\left[-\beta H_0\right]}{\left\{\int \mathrm{d}\Gamma \exp\left[-\beta H_0\right]\right\}} - \frac{\int \mathrm{d}\Gamma A \exp\left[-\beta H_0\right]}{\int \mathrm{d}\Gamma \exp\left[-\beta H_0\right]} \times \frac{\int \mathrm{d}\Gamma \beta D \exp\left[-\beta H_0\right]}{\int \mathrm{d}\Gamma \exp\left[-\beta H_0\right]}$$

Giving: $$\left\langle \frac{\partial A}{\partial \varepsilon} \right\rangle_0 = \beta\left\{\langle AD \rangle_0 - \langle A \rangle_0 \langle D \rangle_0\right\}$$

If we apply this result for c_A: $\left\langle \frac{\partial A}{\partial \varepsilon} \right\rangle_0 = \beta \left\{ \langle AD \rangle_0 - \langle A \rangle_0 \langle D \rangle_0 \right\}$

with $H = H_0 - \varepsilon g_A \left(q - q^* \right)$

$$\Delta c_A = \langle g_A \rangle_\varepsilon - \langle g_A \rangle_0$$

$$\frac{\partial \Delta c_A}{\partial \varepsilon} = \beta \left(\left\langle \left(g_A \right)^2 \right\rangle_0 - \langle g_A \rangle_0^2 \right)$$

Since g_A =0 or 1:
$g_A(x)\, g_A(x) = g_A(x)$

$$= \beta \left(\langle g_A \rangle_0 \left(1 - \langle g_A \rangle_0 \right) \right)$$

$$= \beta \left(\langle c_A \rangle_0 \left(1 - \langle c_A \rangle_0 \right) \right) = \beta \langle c_A \rangle_0 \langle c_B \rangle_0$$

Giving: $$\boxed{\Delta c_A = \beta \langle c_A \rangle_0 \langle c_B \rangle_0 \, \varepsilon}$$

Linear Response theory (dynamic)

Let us now switch off the perturbation at t=0

$$H = H_0 - \varepsilon D \qquad\qquad H = H_0$$ at t>0:

Let us see how the system relaxes to equilibrium (dynamical perturbation)

We take $\langle A \rangle_0 = 0$

$$\left\langle \Delta A(t) \right\rangle = \left\langle A(t) \right\rangle - \left\langle A \right\rangle_0 = \left\langle A(t) \right\rangle$$

Similar as for the static case for small values of ε, we have

$$\left\langle \frac{\partial A(t)}{\partial \varepsilon} \right\rangle_0 = \frac{\int \mathrm{d}\Gamma \beta A(t) D \exp\left[-\beta H_0\right]}{\left\{\int \mathrm{d}\Gamma \exp\left[-\beta H_0\right]\right\}} = \beta \left\langle D(0) A(t) \right\rangle$$

Giving:

$$\left\langle \Delta A(t) \right\rangle = \beta\varepsilon \left\langle D(0) A(t) \right\rangle$$

$$\left\langle \Delta A(t) \right\rangle = \beta\varepsilon \left\langle D(0)\, A(t) \right\rangle$$

If we apply this result to $D = \Delta g_A \quad \text{and} \quad A = \Delta g_A$

We obtain: $\left\langle \Delta c_A(t) \right\rangle = \beta\varepsilon \left\langle \Delta g_A(0) \Delta g_A(t) \right\rangle$

From static perturbation:

$$\beta\varepsilon = \frac{\Delta c_A(0)}{\left\langle c_A \right\rangle \left\langle c_B \right\rangle}$$

$$\left\langle \Delta c_A(t) \right\rangle = \Delta c_A(0) \frac{\left\langle \Delta g_A(0) \Delta g_A(t) \right\rangle}{\left\langle c_A \right\rangle \left\langle c_B \right\rangle}$$

Compare linear response expression with the macroscopic expression

$$\Delta c_A(t) = \Delta c_A(0) \exp\left[-t/\tau \right]$$

Microscopic rate expression

$$\exp\left[-t/\tau\right] = \frac{\left\langle \Delta g_A(0)\,\Delta g_A(t)\right\rangle}{\left\langle c_A\right\rangle\left\langle c_B\right\rangle}$$

Δ has disappeared because of the derivative

Derivative

$$-\frac{1}{\tau}\exp\left[-t/\tau\right] = \frac{\left\langle g_A(0)\,\dot{g}_A(t)\right\rangle}{\left\langle c_A\right\rangle\left\langle c_B\right\rangle} = -\frac{\left\langle \dot{g}_A(0)\,g_A(t)\right\rangle}{\left\langle c_A\right\rangle\left\langle c_B\right\rangle}$$

Stationary (t is arbitrary, only depends on τ)

$$\frac{d}{dt}\left\langle A(t)\,B(t+\tau)\right\rangle = 0$$

$$\left\langle A(t)\,\dot{B}(t+\tau)\right\rangle + \left\langle \dot{A}(t)\,B(t+\tau)\right\rangle = 0$$

$$\left\langle A(0)\,\dot{B}(\tau)\right\rangle = -\left\langle \dot{A}(0)\,B(\tau)\right\rangle$$

We have
$$\frac{1}{\tau}\exp\left[-t/\tau\right]=\frac{\left\langle \dot{g}_A\left(0\right)g_A\left(t\right)\right\rangle}{\left\langle c_A\right\rangle\left\langle c_B\right\rangle}$$

Using
$$\tau=k_{A\to B}^{-1}\left(1+\left\langle c_A\right\rangle/\left\langle c_B\right\rangle\right)^{-1}=\frac{\left\langle c_B\right\rangle}{k_{A\to B}}$$

For sufficiently short t, we obtain
$$\boxed{k_{A\to B}\left(t\right)=\frac{\left\langle \dot{g}_A\left(0\right)g_A\left(t\right)\right\rangle}{\left\langle c_A\right\rangle}}$$

Using the definition of g_A we can write
$$\dot{g}_A\left(q-q^*\right)=\dot{q}\frac{\partial g_A\left(q-q^*\right)}{\partial q}=-\dot{q}\frac{\partial g_B\left(q-q^*\right)}{\partial q}$$

$$k_{A\to B}\left(t\right)=\frac{\left\langle \dot{q}\left(0\right)\frac{\partial g_B\left(q\left(0\right)-q^*\right)}{\partial q}g_B\left(t\right)\right\rangle}{\left\langle c_A\right\rangle}$$

We now have an expression that relates the macroscopic reaction rate to microscopic properties

$$k_{A\to B}(t)=\frac{\left\langle \dot{q}(0)\dfrac{\partial g_B\left(q(0)-q^*\right)}{\partial q}\,g_B(t)\right\rangle}{\langle c_A\rangle}$$

Let us look at the different terms in this equation

$$g_B(t)=\theta\left(q(t))-q^*\right)$$

Only when the system is in the product state we get a contribution to the ensemble average

$$\frac{\partial g_B\left(q(0)-q^*\right)}{\partial q}=\frac{\partial\Theta\left(q(0)-q^*\right)}{\partial q}$$

$$=\delta\left(q(0)-q^*\right)$$

Only when the system starts at the transition state, we get a contribution to the ensemble average

$$\dot{q}(0)$$

Velocity at t=0

$$\langle c_A\rangle=\left\langle\Theta\left(q^*-q\right)\right\rangle$$

Concentration of A

$$\boxed{k_{A\to B}(t)=\frac{\left\langle \dot{q}(0)\delta\left(q(0)-q^*\right)\theta\left(q(t)-q^*\right)\right\rangle}{\left\langle\theta\left(q^*-q\right)\right\rangle}}$$

Transition state theory

$$k_{A\to B}(t) = \frac{\left\langle \dot{q}(0)\delta\left(q(0)-q^*\right)\theta\left(q(t)-q^*\right)\right\rangle}{\left\langle \theta\left(q^*-q\right)\right\rangle}$$

Let us consider the limit: t → 0^+

$$\lim_{t\to 0^+} = \theta\left(q(t)-q^*\right) = \theta\left(\dot{q}(t)\right)$$

Contribution for positive velocities

This gives:

$$k_{A\to B}^{TST} = \frac{\left\langle \dot{q}(0)\delta\left(q(0)-q^*\right)\theta\left(\dot{q}\right)\right\rangle}{\left\langle \theta\left(q^*-q\right)\right\rangle}$$

Eyring's transition state theory

Transition state theory

$$k_{A\to B}(t) = \frac{\left\langle \dot{q}(0)\,\delta\big(q(0)-q^*\big)\,\theta\big(q(t)-q^*\big)\right\rangle}{\left\langle \theta\big(q^*-q\big)\right\rangle}$$

We can rewrite this expression as a product by inserting 1

$$k_{A\to B}(t) = \frac{\left\langle \dot{q}(0)\,\delta\big(q(0)-q^*\big)\,\theta\big(q(t)-q^*\big)\right\rangle}{\left\langle \delta\big(q(0)-q^*\big)\right\rangle} \times \frac{\left\langle \delta\big(q(0)-q^*\big)\right\rangle}{\left\langle \theta\big(q^*-q\big)\right\rangle}$$

Conditional “probability” to find a particle on the top of the barrier with a positive velocity

Ratio of probabilities to find particle on top of the barrier and in the state A

$$k_{A\to B}(t) = \left\langle \dot{q}(0)\,\theta\big(q(t)-q^*\big)\right\rangle_{q=q^*} \times \frac{\left\langle \delta\big(q(0)-q^*\big)\right\rangle}{\left\langle \theta\big(q^*-q\big)\right\rangle}$$

Free energy barrier

$$\frac{\left\langle \delta\left(q(0)-q^*\right)\right\rangle}{\left\langle \theta\left(q^*-q\right)\right\rangle}$$

Ratio of the probabilities to find a particle on top of the barrier and in the state A

Probability to be on top of the barrier:

$$\left\langle \delta\left(q^*-q\right)\right\rangle = C\int dq\,\delta\left(q-q^*\right)\exp\left(-\beta F(q)\right) = C\exp\left(-\beta F\left(q^*\right)\right)$$

Probability to be in state A:

$$\left\langle \Theta\left(q^*-q\right)\right\rangle = C\int dq\,\Theta\left(q-q^*\right)\exp\left(-\beta F(q)\right) = C\int_{q<q^*} dq\exp\left(-\beta F(q)\right)$$

This gives:

$$\frac{\left\langle \delta\left(q(0)-q^*\right)\right\rangle}{\left\langle \theta\left(q^*-q\right)\right\rangle} = \frac{\exp\left(-\beta F\left(q^*\right)\right)}{\int_{q<q^*} dq\exp\left(-\beta F(q)\right)}$$

We need to determine the free energy as a function of the order parameter

$$\left\langle \dot{q}(0)\,\theta\big(q(t)-q^*\big)\right\rangle_{q=q^*}$$

Conditional "probability" to find a particle on the top of the barrier with a positive velocity

$$\dot{q}(0)$$

Assume that on top of the barrier the particle is in equilibrium: use Maxwell-Boltzmann distribution to generate this velocity

$$\dot{q}(0)\,\theta\big(q(t)-q^*\big)$$

Only particles with a positive velocity end up in the product state. We assume that once in the product state they stay there.

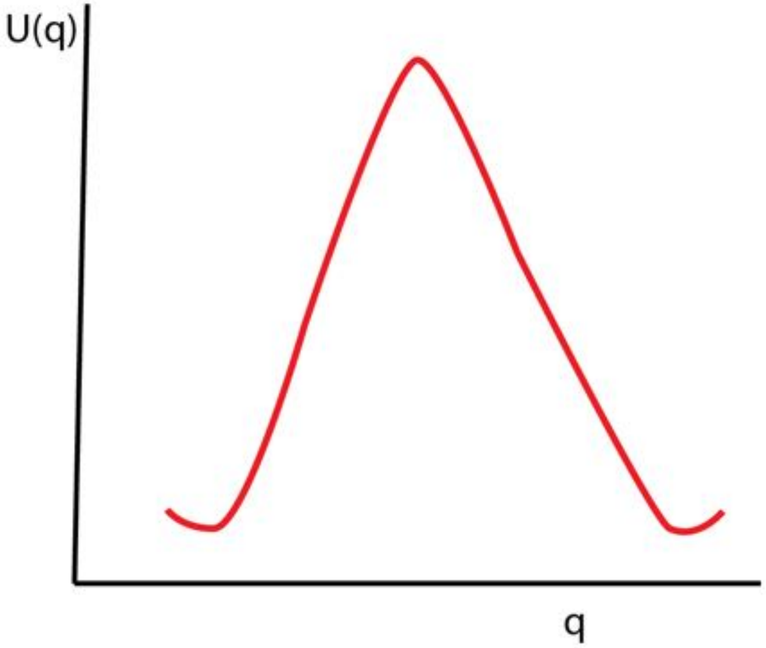
$$\lim_{t\to 0^+}\dot{q}(0)\,\theta\big(q(t)-q^*\big)=\dot{q}(0)\,\theta\big(\dot{q}(0)\big)=0.5\left|\dot{q}(0)\right|$$

$$k_{A\to B}^{TST}=\lim_{t\to 0^+}\left\langle \dot{q}(0)\,\theta\big(q(t)-q^*\big)\right\rangle_{q=q^*}\times\frac{\left\langle \delta\big(q(0)-q^*\big)\right\rangle}{\left\langle \theta\big(q^*-q\big)\right\rangle}$$

$$\boxed{k_{A\to B}^{TST}=0.5\left|\dot{q}(0)\right|\frac{\exp\big(-\beta F(q^*)\big)}{\int_{q<q^*} dq\,\exp\big(-\beta F(q)\big)}}$$

Eyring's TST

1-D ideal gas particle on a hill



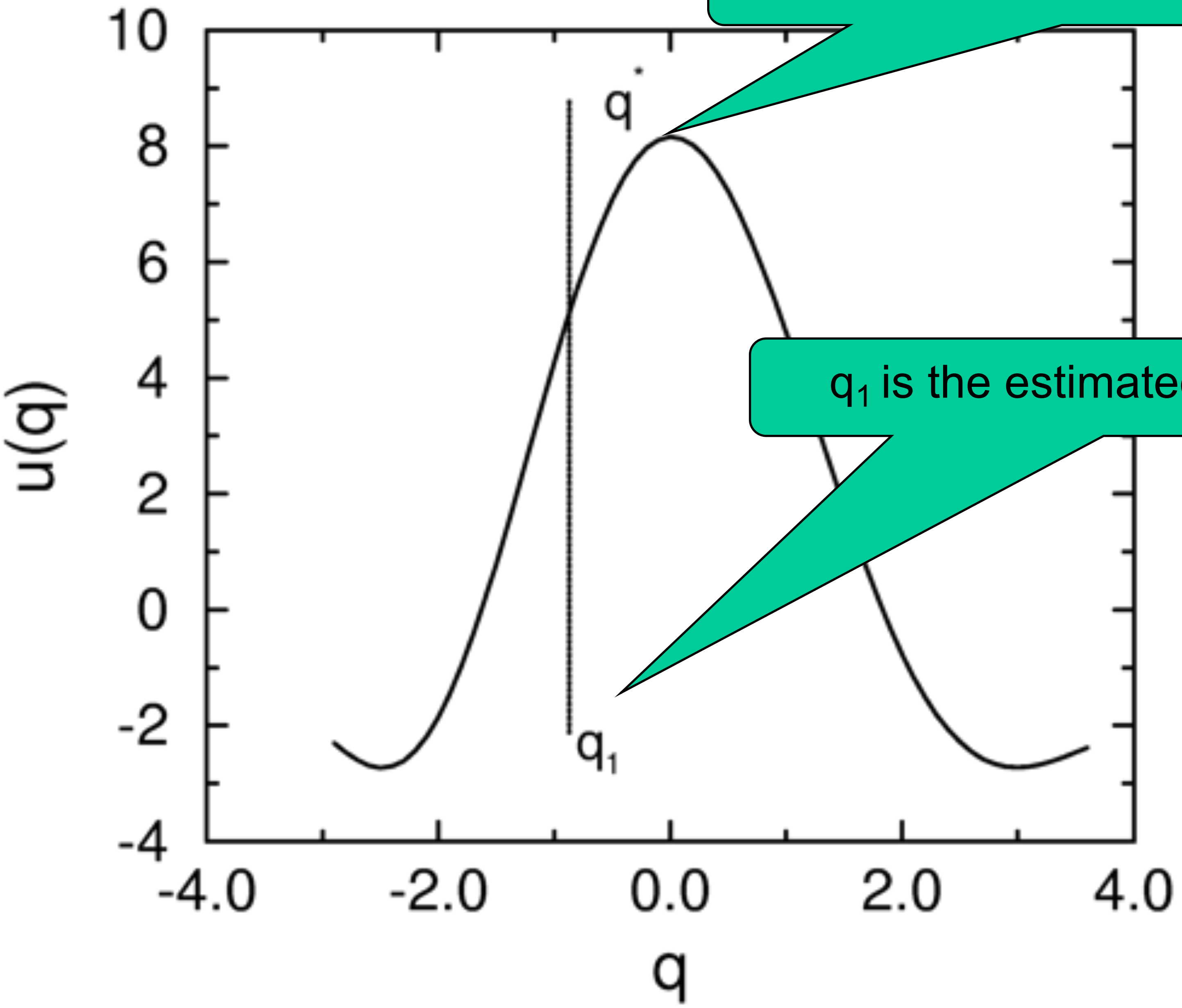
$$k_{A\to B}^{TST} = 0.5\left|\dot{q}(0)\right| \frac{\exp\left(-\beta F\left(q^*\right)\right)}{\int\limits_{q<q^*} dq \exp\left(-\beta F\left(q\right)\right)}$$

Maxwell-Boltzmann: $\left|\dot{q}(0)\right| = \sqrt{\dfrac{2k_B T}{\pi m}}$

This gives for the hopping rate

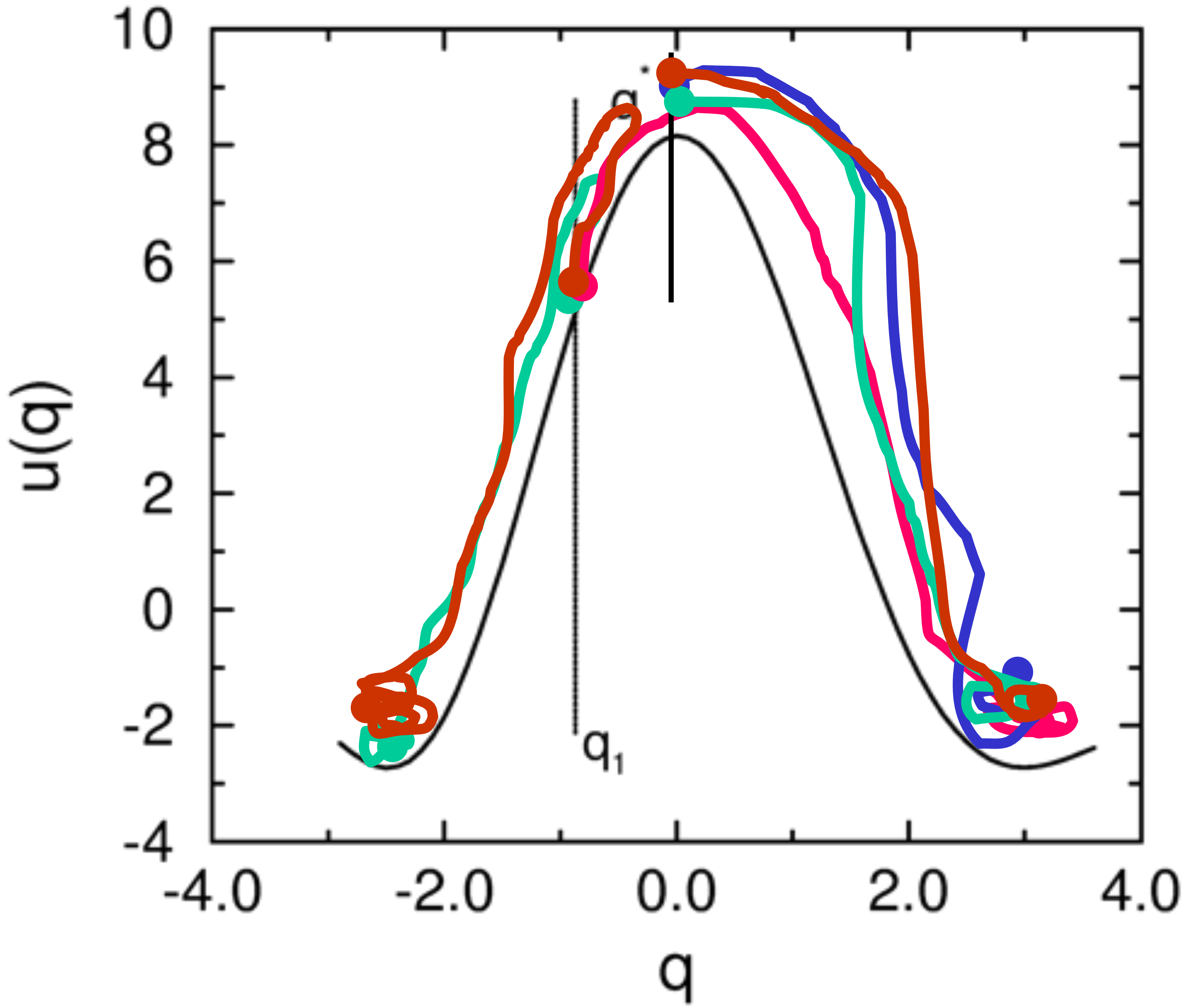
$$k_{A\to B}^{TST} = \sqrt{\frac{k_B T}{2\pi m}} \frac{\exp\left(-\beta U\left(q^*\right)\right)}{\int\limits_{q<q^*} dq \exp\left(-\beta U\left(q\right)\right)}$$

Ideal gas particle on a not-so-ideal hill



q^* is the true transition state

q_1 is the estimated transition state

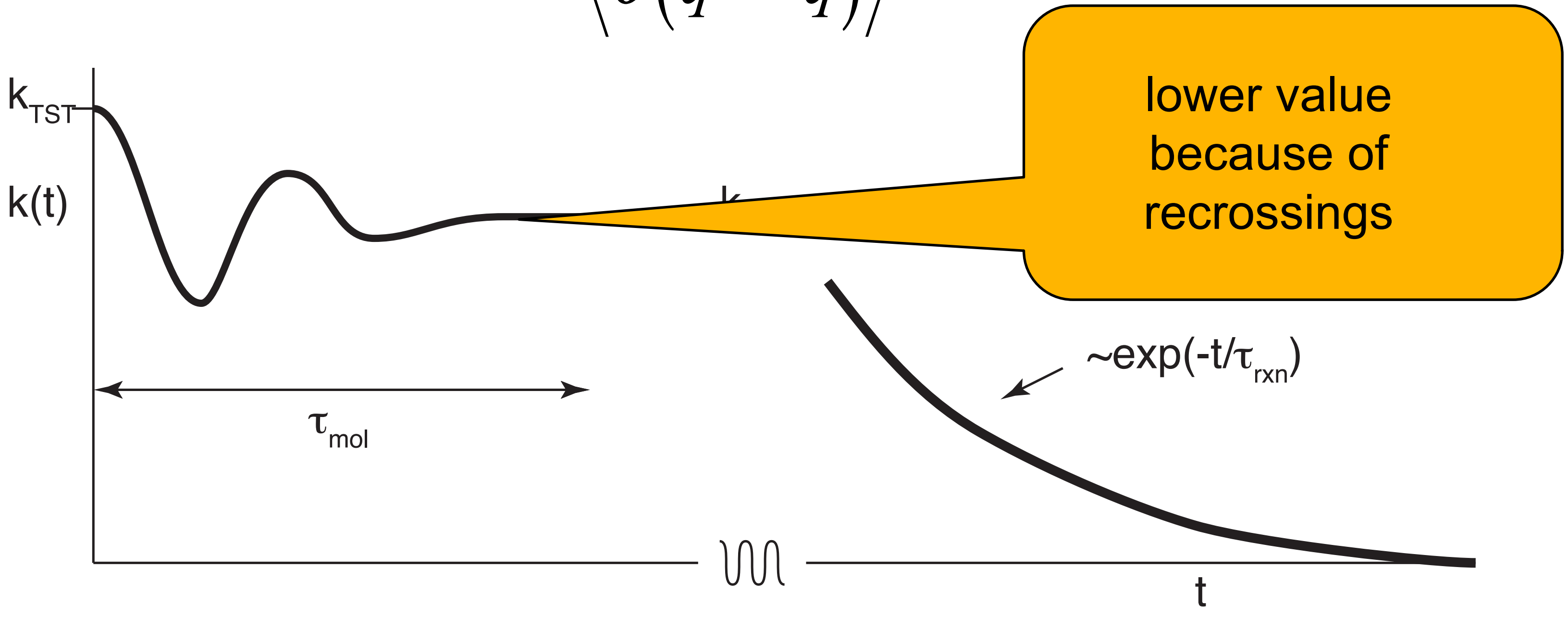


For this case transition state theory will overestimate the hopping rate

Transition state theory

- One has to know the free energy accurately (MC/MD)
- Gives only an upper bound to the reaction rate
- Assumptions underlying transition theory should hold: *no recrossings*

$$k_{A\to B}(t)=\frac{\left\langle \dot{q}(0)\,\delta\left(q(0)-q^*\right)\theta\left(q(t)-q^*\right)\right\rangle}{\left\langle \theta\left(q^*-q\right)\right\rangle}$$



lower value because of recrossings

Outline

- Part 1
  - Rare event and reaction kinetics
  - Linear Response theory
  - Transition state theory
  - **Free energy methods**
  - Bennet Chandler approach
  - Example zeolites

- Part 2
  - Two ended methods
  - Transition path sampling
  - Rate constants
  - Reaction coordinate analysis
  - Application to crystallization
  - Path metadynamics

Free energy barriers in complex systems

- Straightforward MD or MC and then use $\beta F(q) = -\ln\left\langle \delta\left(q(r) - q\right)\right\rangle$ is highly inefficient for high barriers

- Many "tricks" have been proposed to overcome and sample barriers
  - Temperature enhanced sampling: simulated tempering, parallel tempering, Temperature accelerated molecular dynamics …)
  - Constraint dynamics: thermodynamic integration, blue moon sampling....
  - Flat histogram sampling: umbrella sampling, hyperdynamics,....
  - history dependent search: Wang-Landau, adaptive biasing force, metadynamics,…
  - non-equilibrium methods: steered MD, targeted MD,....

  - trajectory-based methods: nudged elastic band, action minimization, string method, transition path sampling, forward flux sampling,....

Free energy barriers

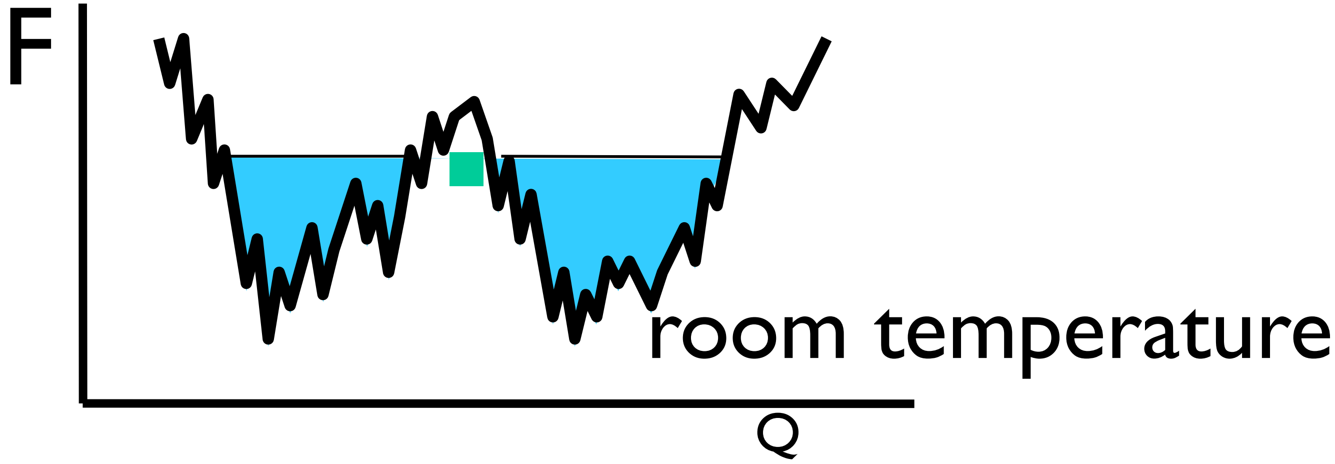
- Replica exchange

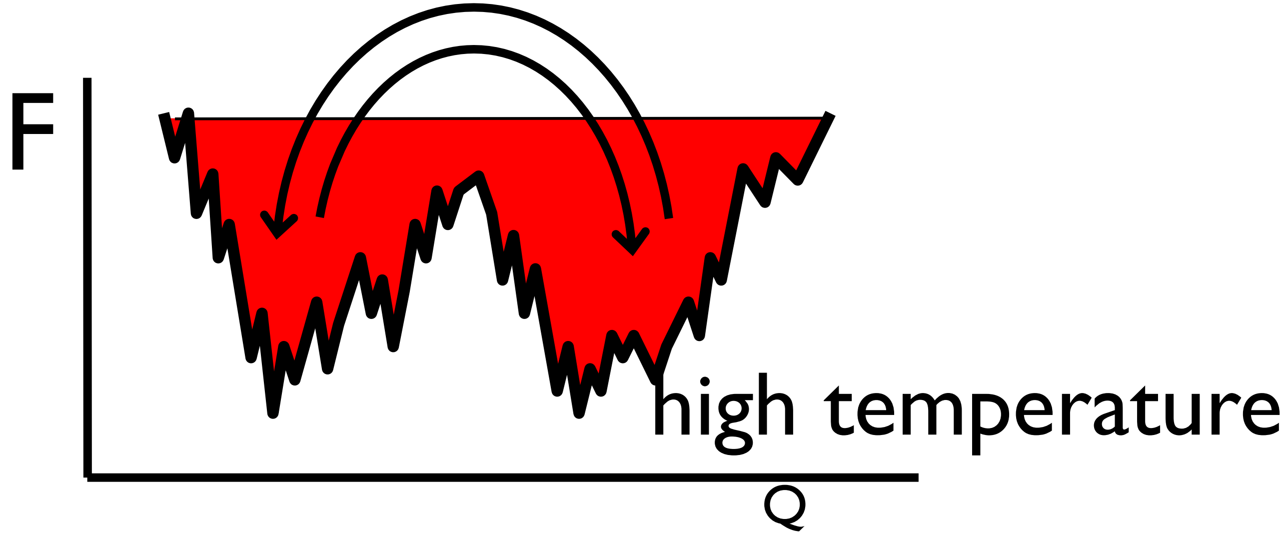
- Thermodynamic integration

- Umbrella sampling

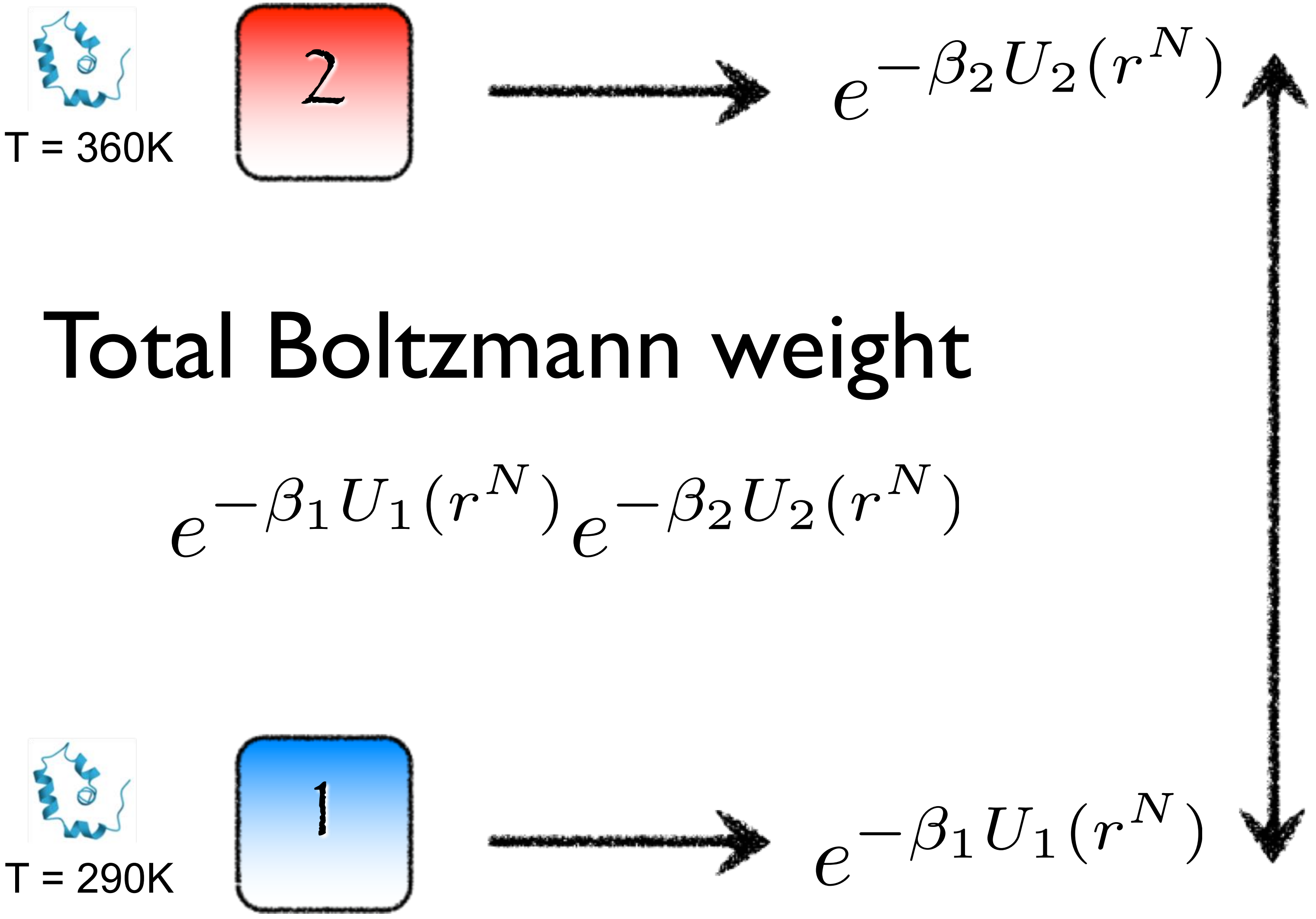
- Metadynamics

Replica exchange/parallel tempering





Two replicas



T = 360K

2

$$e^{-\beta_2 U_2(r^N)}$$

Total Boltzmann weight

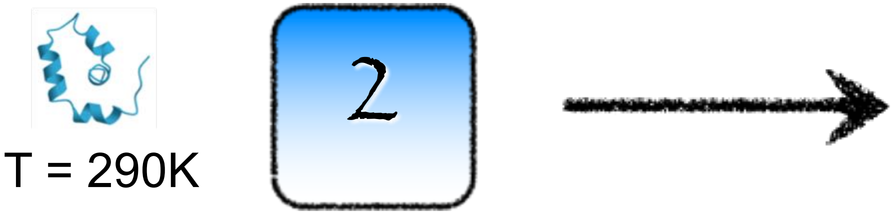
$$e^{-\beta_1 U_1(r^N)} e^{-\beta_2 U_2(r^N)}$$

T = 290K

1

$$e^{-\beta_1 U_1(r^N)}$$

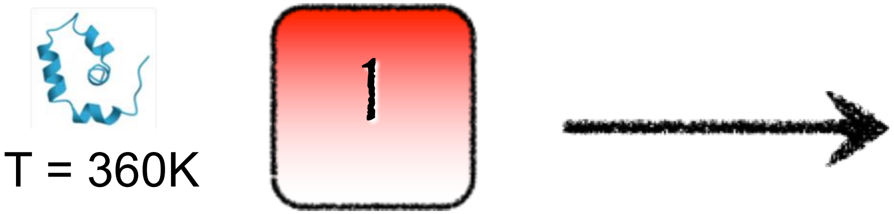
Switching temperatures



$$e^{-\beta_1 U_2(r^N)}$$

Total Boltzmann weight

$$e^{-\beta_1 U_2(r^N)} e^{-\beta_2 U_1(r^N)}$$



$$e^{-\beta_2 U_1(r^N)}$$

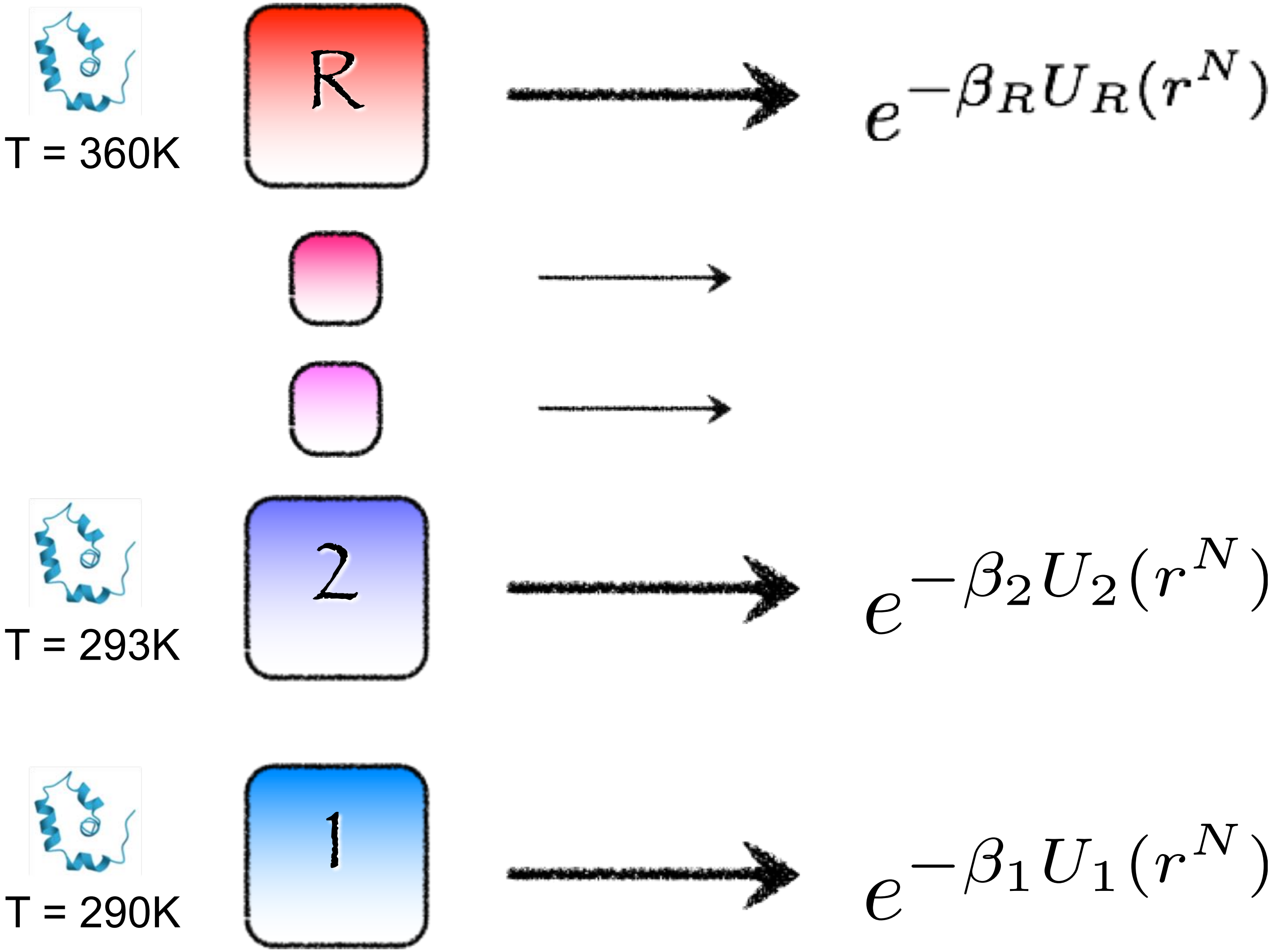
The ratio of the new Boltzmann factor over the old one is:

$$\frac{\mathcal{N}(n)}{\mathcal{N}(o)} = e^{(\beta_2 - \beta_1)[U_2(r^N) - U_2(r^N)]}$$

the rule for switching temperatures should obey detailed balance
Metropolis Monte Carlo scheme

$$\mathrm{acc}(1 \leftrightarrow 2) = \min\left(1, e^{(\beta_2 - \beta_1)[U_2(r^N) - U_1(r^N)]}\right)$$

Set of replicas



T = 360K

$$e^{-\beta_R U_R(r^N)}$$

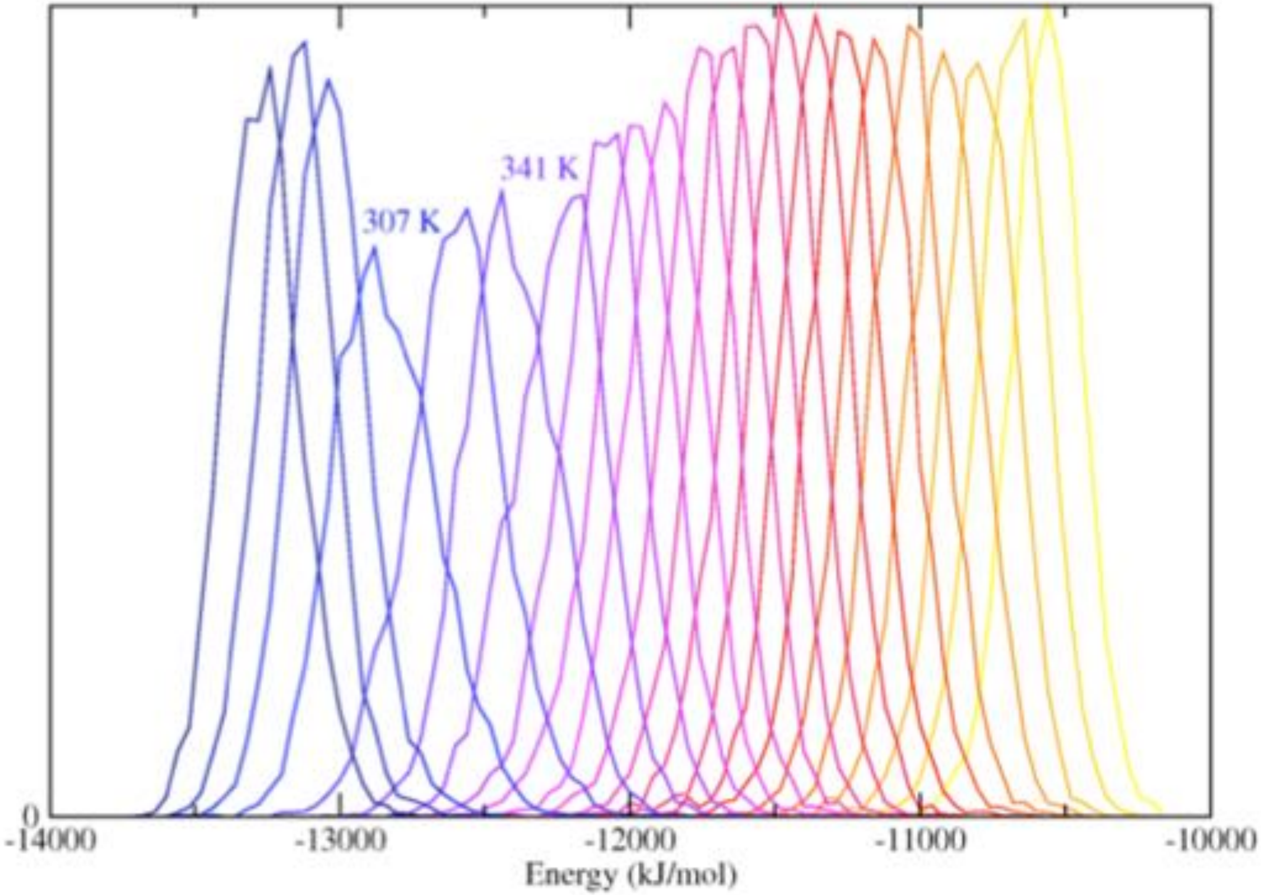
T = 293K

$$e^{-\beta_2 U_2(r^N)}$$

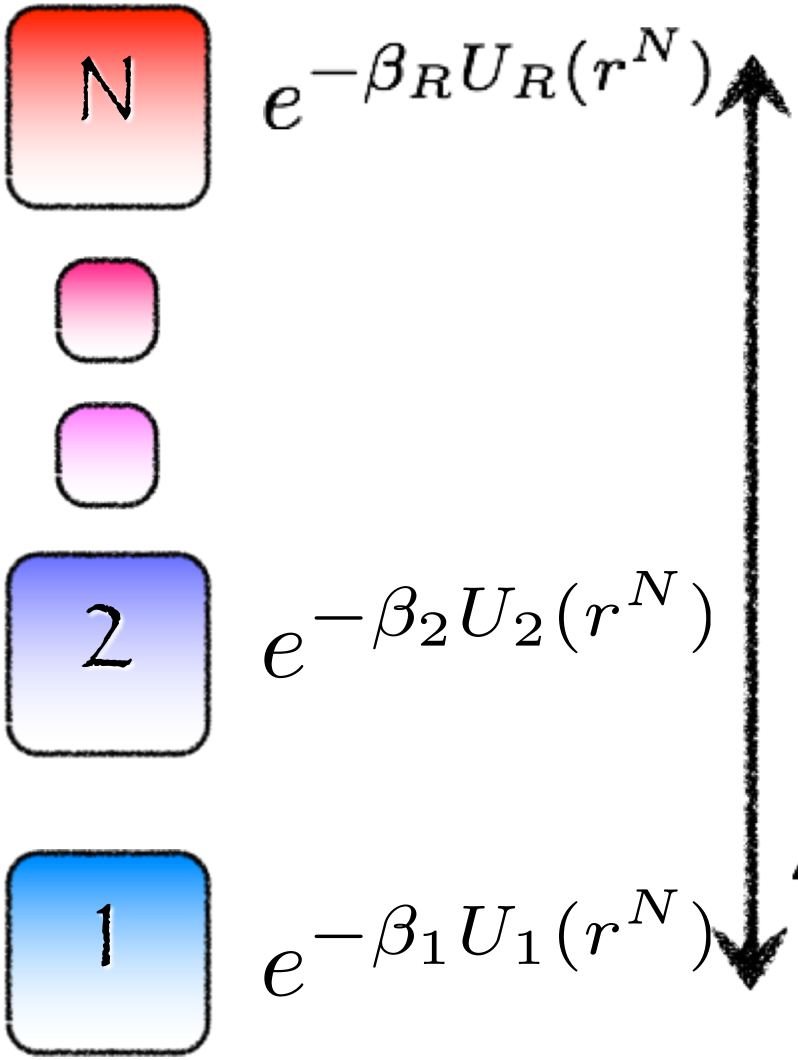
T = 290K

$$e^{-\beta_1 U_1(r^N)}$$

Overlap in potential energy



Replica Exchange MD (REMD)



$$e^{-\beta_R U_R(r^N)}$$

$$e^{-\beta_2 U_2(r^N)}$$

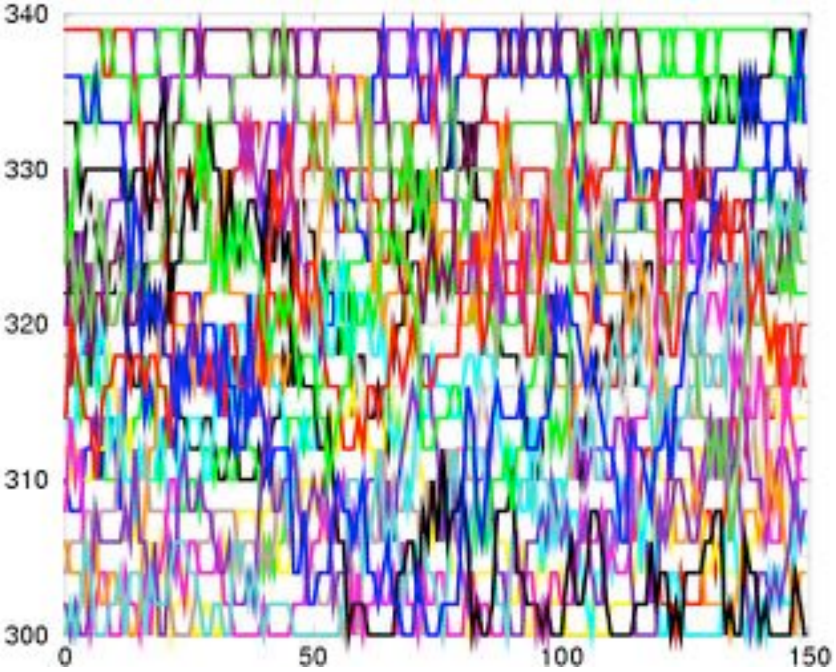
$$e^{-\beta_1 U_1(r^N)}$$

$$P_{acc} = min[1, exp(-\Delta_{ij})]$$

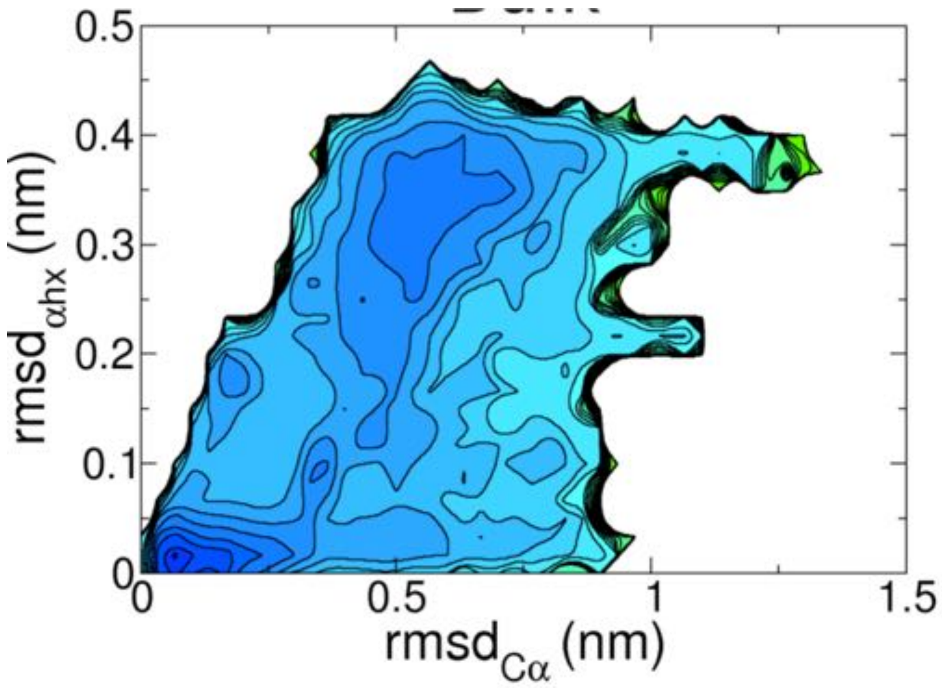
$$\Delta_{ij} = (\beta_i - \beta_j)(U_i(r^N) - U_j(r^N))$$

Hansmann *Chem Phys Lett* 1997
Sugita & Okamoto *Chem Phys Lett* 1999

Replica Exchange



Exchange as a function of time.



Free energy follows from

$$\beta F(x,y) = -\ln P(x,y)$$

Advantage: no order parameters needed

Disadvantage: convergence of free energy landscape can be still slow, especially around phase transition: many replicas needed.

Free energy barriers

- Replica exchange
- **Thermodynamic integration**
- Umbrella sampling
- Metadynamics

Thermodynamic integration

- The free energy follows from the derivative

$$\beta F(\lambda_1) - \beta F(\lambda_0) = \int_{\lambda_0}^{\lambda 1} \left(\frac{\partial F}{\partial \lambda} \right) d\lambda$$

- The derivative of the free energy is known as the mean force

$$\left(\frac{\partial F}{\partial \lambda} \right) = \left\langle \frac{dU}{\partial \lambda} \right\rangle \equiv \langle f(\lambda) \rangle$$

- compute the force f at λ directly or by adding a constraint to the Lagrangian

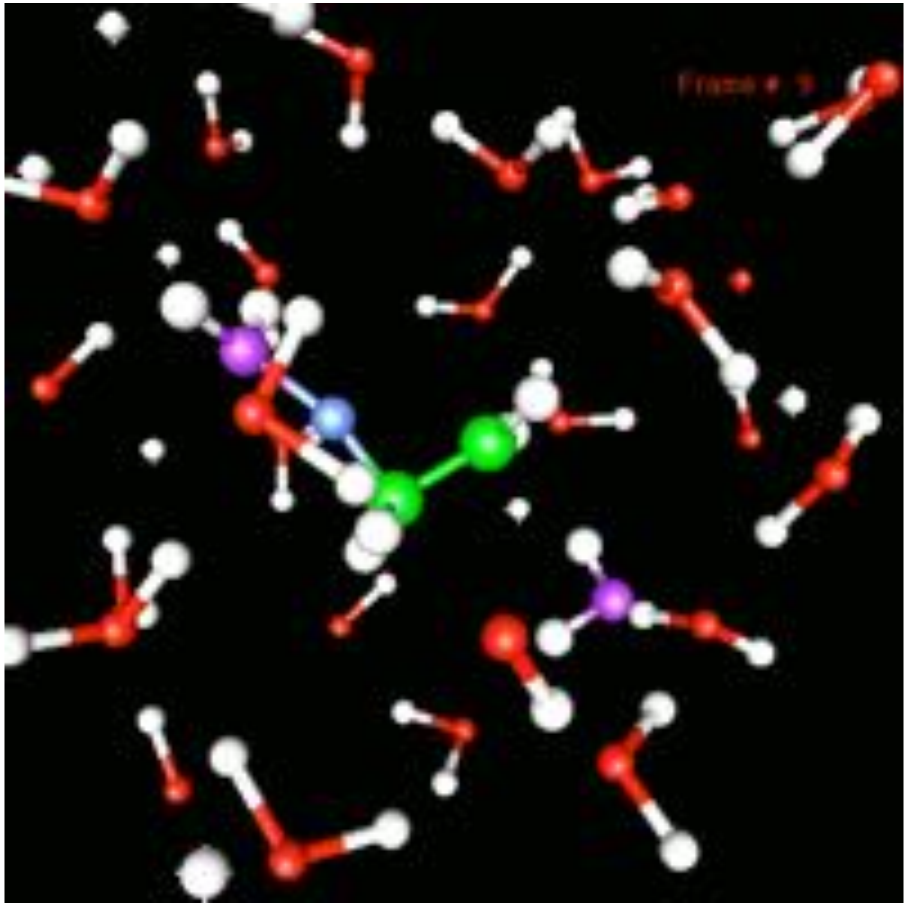
$$\mathcal{L}_c(\dot{r}^N, r^N) = \mathcal{L}(\dot{r}^N, r^N) - g(\sigma(r^N) - \lambda)$$

constraint force

- the constraint force follows from the Lagrange multiplier

$$\langle f(\lambda) \rangle = \langle g \rangle$$

Example: Alkene hydration



SYSTEM

$32\ H_2O\ + H^+\ + C_2H_4$

T=300K

REACTION COORDINATE Q

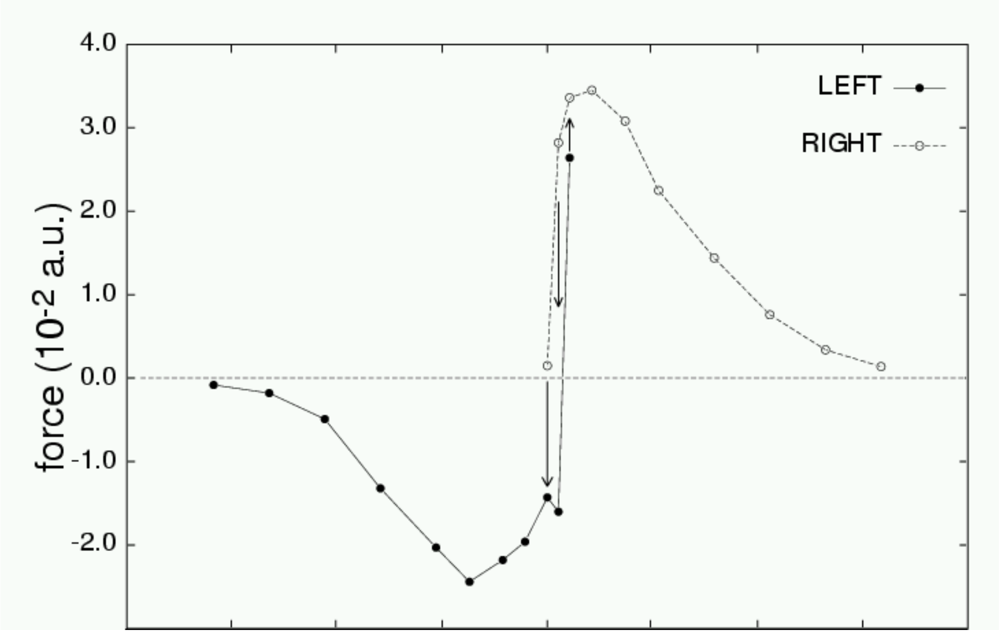
$Q = R_{OH} - R_{HC}$

$C_2H_4 + H_2O \qquad CH_3CH_20H$

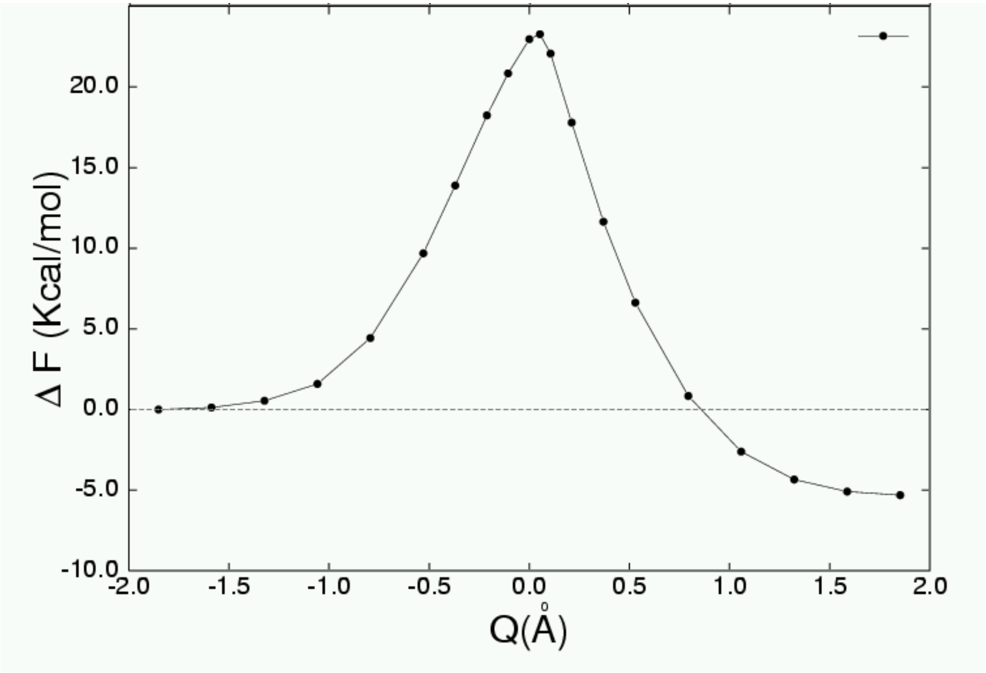
T.Van Erp, E-J Meijer ,
Angew. Chem, 43, 1660 (2004).

Example: Alkene hydration

CONSTRAINT FORCE



FREE ENERGY PROFILE



	kcal/mol
CPMD-BLYP	23
Exp: Gas Phase	50-100
MP2: Gas Phase	58
Exp: Low Density Acid Solution	33
BLYP: Gas Phase + Acid	24

Free energy barriers

- Replica exchange
- Thermodynamic integration
- **Umbrella sampling**
- Metadynamics

Umbrella sampling

The regular distribution of an order parameter q is

$$P(q) = \langle [\delta(q - q(x))] \rangle = \frac{\int dx \exp\left[-\beta U(x)\right] \delta(q - q(x))]}{\int dx \exp\left[-\beta U(x)\right])}$$

Applying a bias potential V_{bs} (q) gives for the biased distribution

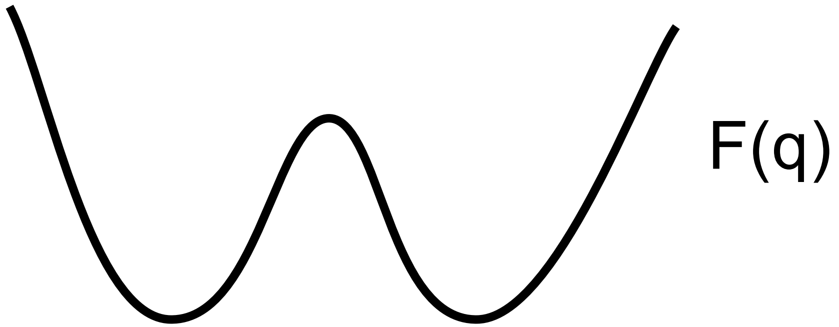
$$P_{bs}(q) = \frac{\int dx \exp\left[-\beta U(x) - \beta V_{bs}(q(x)\right] \delta(q - q(x))]}{\int dx \exp\left[-\beta U(x) - \beta V_{bs}(q(x)\right])}$$

The free energy can be extracted from P_{bs}(q) by

$$\beta F(q) = \text{-}\ln P_{bs}(q) - \beta V_{bs}(q) + const$$

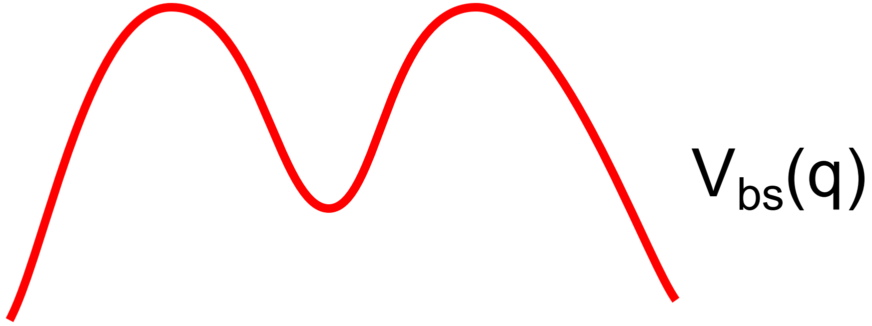
Flat sampling

- Consider a free energy landscape with two minima

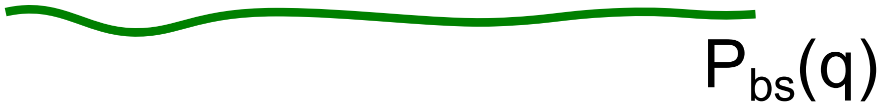


- taking a biasing potential

$$V_{bs}(q) = -F(q)$$



- results in a flat histogram



- This turns out to effectively sample the entire free energy barrier

Umbrella sampling

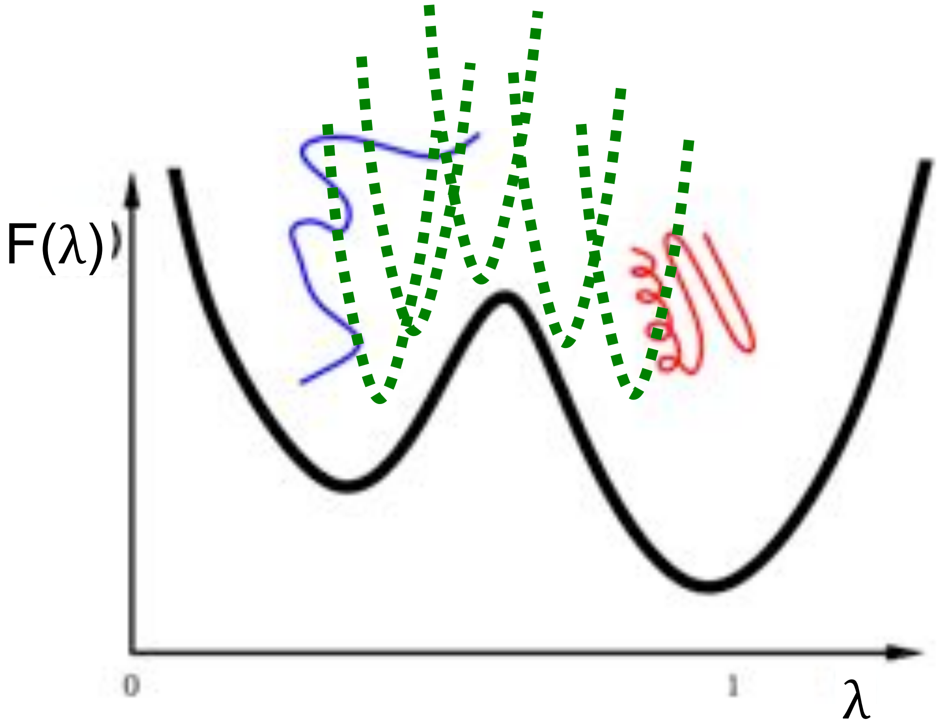
Biasing potential can take any functional form to force system into unlikely region

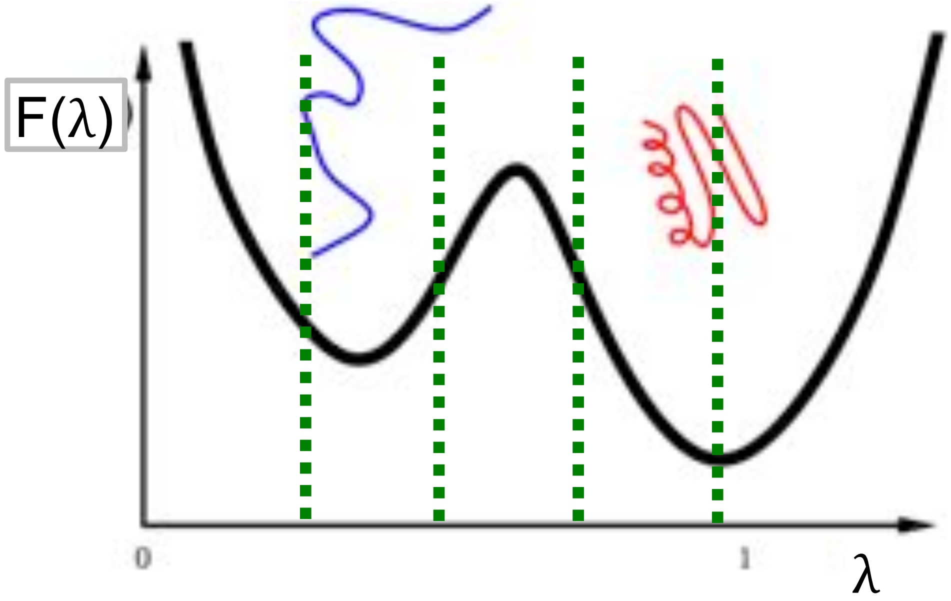
quadratic bias

$$V_{bs}^{i}(\lambda(x)) = c(\lambda(x) - \lambda_i)^2$$

hard window bias

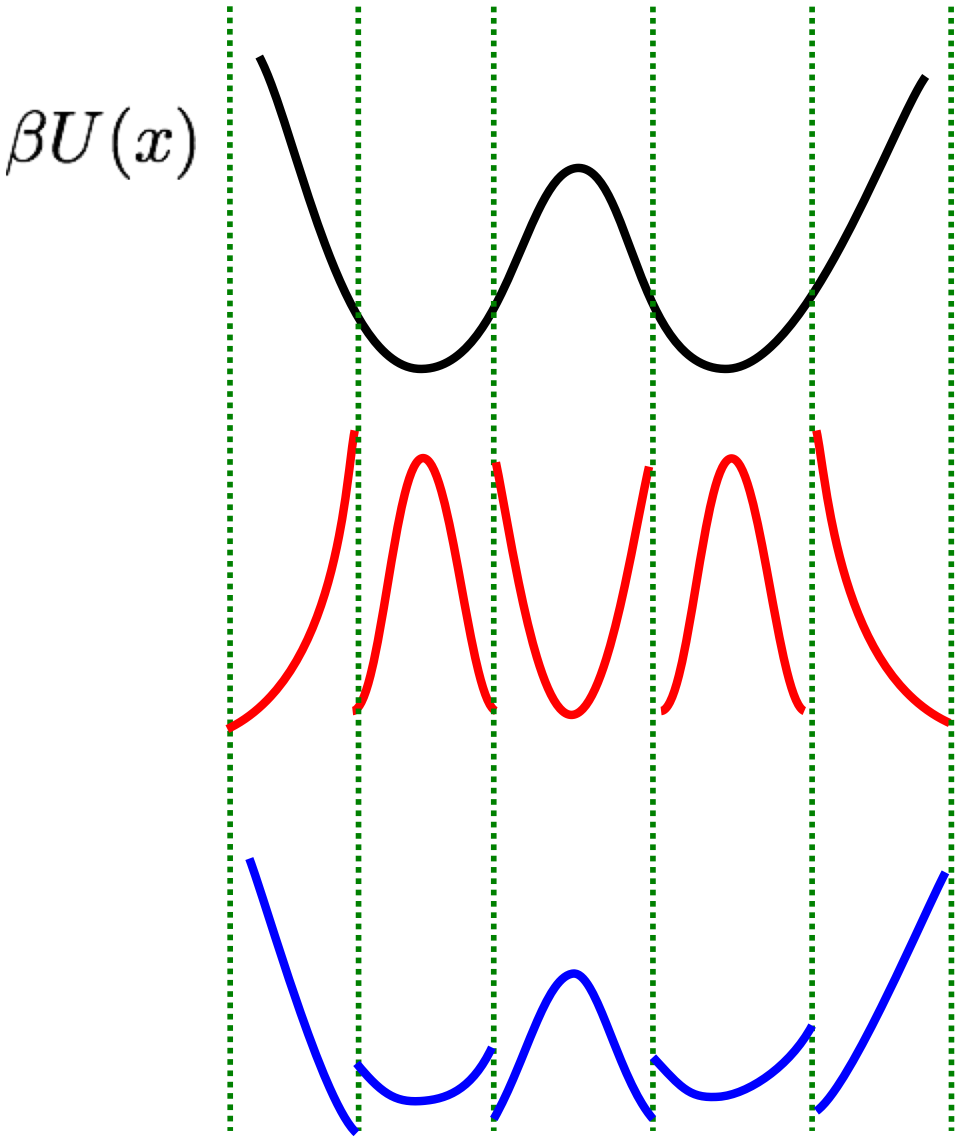
$$V_{bs}^{i}(\lambda(x)) = \begin{cases} 0 & \lambda_i < \lambda(x) < \lambda_{i+1} \\ \infty & \text{otherwise} \end{cases}$$





Histograms

Suppose we perform a hard window simulation



$$V_{bs}^{i}(\lambda(x)) = \begin{cases} 0 & \lambda_i < \lambda(x) < \lambda_{i+1} \\ \infty & \text{otherwise} \end{cases}$$

$$P_{bs}(x)$$

$$\beta F^{i}(x) = -\ln P_{bs}^{i}(x) - \beta V_{bs}^{i}(x)$$

Weighted Histogram Analysis Method

Ferrenberg & Swendsen 1986, Kumar et al 1992

Joins multiple overlapping histograms using an maximum likelihood criterion

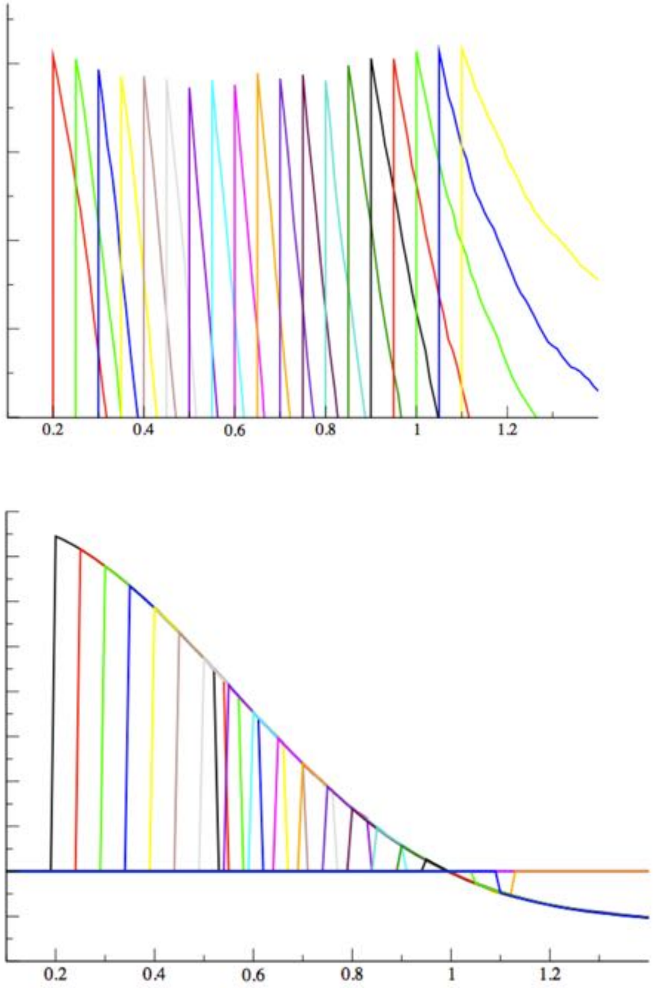
For N_{sims} histograms $n_i(x)$ the best estimate for the joint histogram is

$$p(x) = \frac{\sum_{i=1}^{N_{\text{sims}}} n_i(x)}{\sum_{i=1}^{N_{\text{sims}}} e^{-\beta V_{\text{bs}}^i(x)} N_i / Z_i}$$

where N_i is the total number of measurements in the histogram and Z_i is a “partition function” determined by

$$Z_i = \sum_{x\text{bins}} p(x) e^{-\beta V_{\text{bs}}^i(x)}$$

the two equations have to be solved iteratively



Equivalent to MBAR (see Frenkel’s lecture

Free energy barriers

- Replica exchange
- Thermodynamic integration
- Umbrella sampling
- **Metadynamics**

Metadynamics

- method to obtain free energy in a single simulation
- similar idea as Wang Landau sampling: add history dependent biasing potential to forcefield

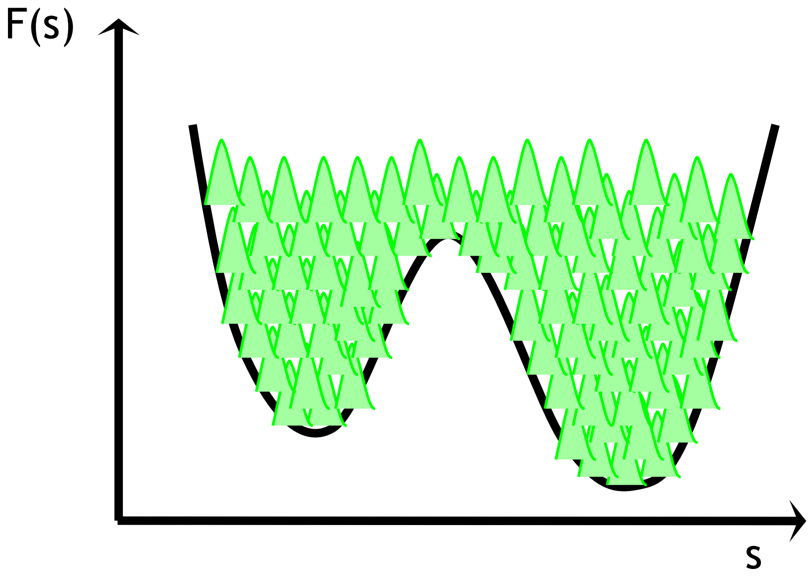
$$V(s;t) = w \sum_{t'<t} e^{-\frac{(s-s(x(t'))^2}{2\sigma}}$$

Laio and Parrinello, PNAS (2002)

- s = predefined order parameters
- w = height of hills
- σ = width of gaussians

- w is reduced every cycle

$$F(s) = -\lim_{t\to\infty} V(s;t)$$



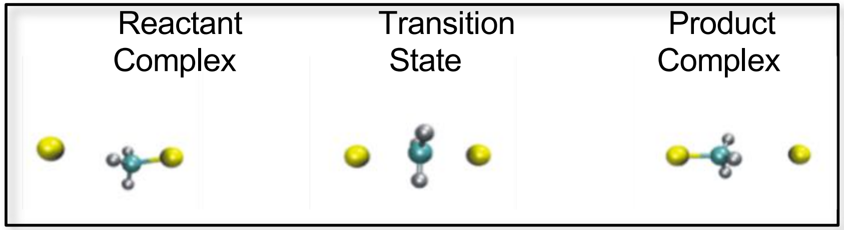
- more controlled version: well tempered MetaD

Barducci, Bussi, Parrinello, PRL, (2008).

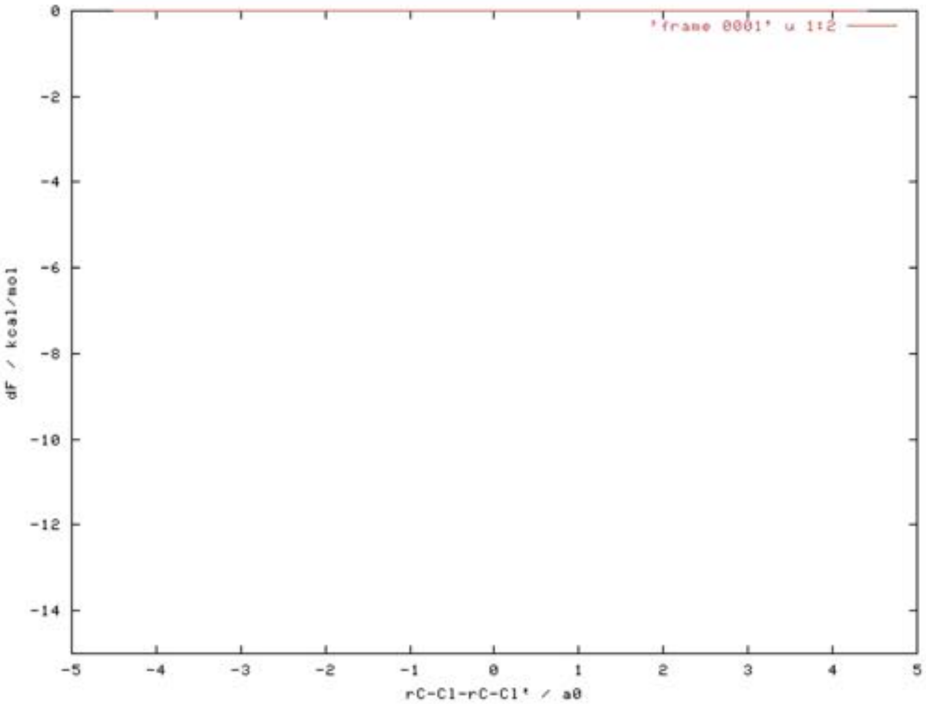
Link to bernds animation

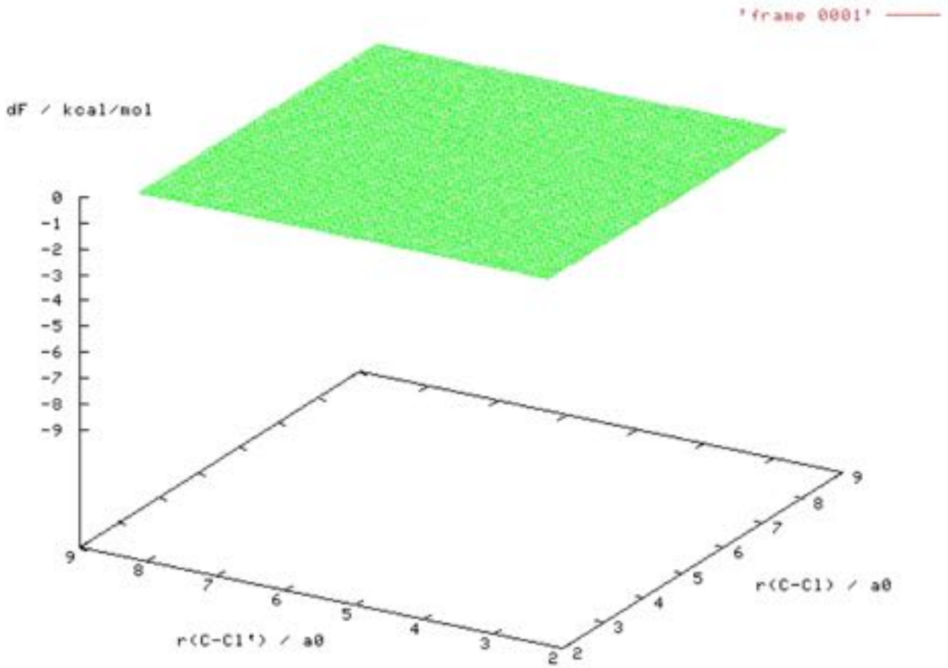
S_N2 reaction between Cl^- and CH_3Cl

$S(R) = r_{C\text{-}Cl} - r_{C\text{-}Cl^-}$



$S_1(R) = r_{C\text{-}Cl}$
$S_2(R) = r_{C\text{-}Cl^-}$





Meta-dynamics can relax the requirement of choosing a good reaction coordinate

Bernd Ensing, Alessandro Laio, Michele Parrinello and Michael L. Klein, *J. Phys. Chem. B* **109** (2005), 6676-6687

Outline

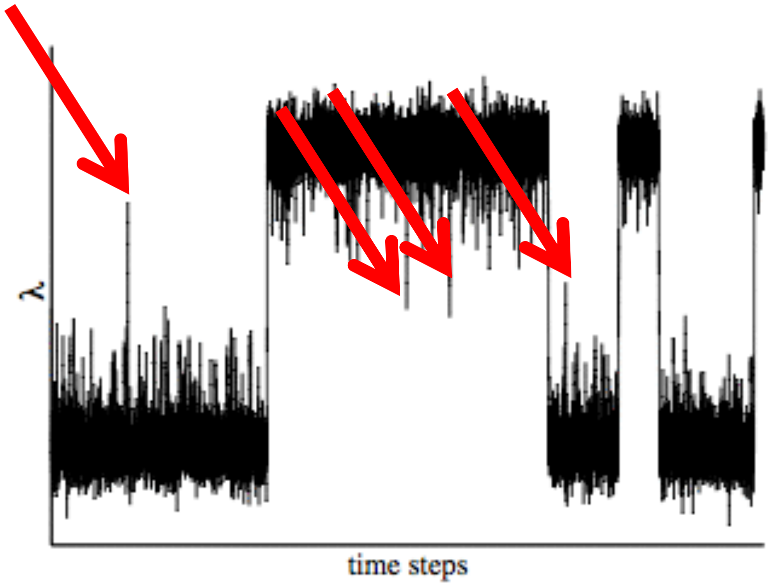
- Part 1
  - Rare event and reaction kinetics
  - Linear Response theory
  - Transition state theory
  - Free energy methods
  - **Bennet Chandler approach**
  - Example zeolites

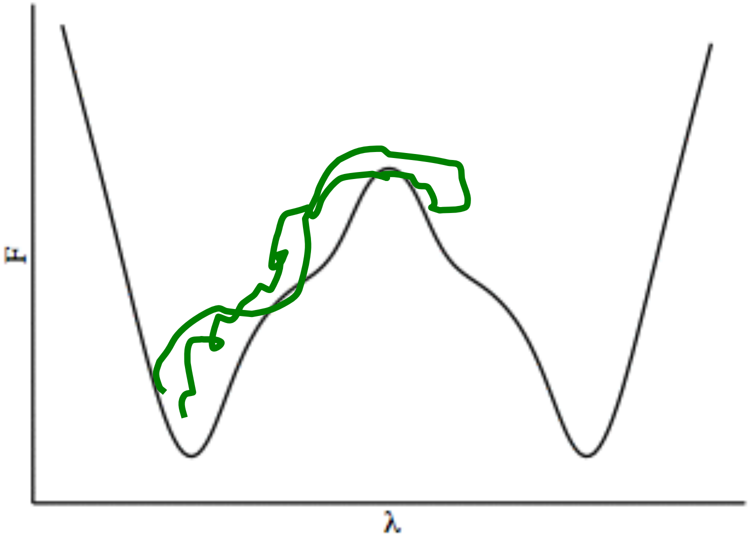
- Part 2
  - Two ended methods
  - Transition path sampling
  - Rate constants
  - Reaction coordinate analysis
  - Application to crystallization
  - Path metadynamics

Problem with TST

There are recrossings that cause overestimation of the rate constant

trajectories that seem to overcome the barrier but in fact bounce back





Bennett-Chandler approach

$$k_{A\to B}(t) = \frac{\left\langle \dot{q}(0)\,\delta\left(q(0)-q^*\right)\theta\left(q(t)-q^*\right)\right\rangle}{\left\langle \theta\left(q^*-q\right)\right\rangle}$$

$$k_{A\to B}(t) = \frac{\left\langle \dot{q}(0)\,\delta\left(q(0)-q^*\right)\theta\left(q(t)-q^*\right)\right\rangle}{\left\langle \delta\left(q(0)-q^*\right)\right\rangle} \times \frac{\left\langle \delta\left(q(0)-q^*\right)\right\rangle}{\left\langle \theta\left(q^*-q\right)\right\rangle}$$

Computational scheme:

1. Determine the probability from the free energy using MC or MD, e.g. by umbrella sampling, thermodynamic integration or other free energy methods
2. Compute the conditional average from a MD simulation

Bennett-Chandler approach

$$k_{A\to B}^{TST}(t)=\frac{\langle \dot{q}(0)\delta(q(0)-q_1)\theta(\dot{q})\rangle}{\langle \delta(q(0)-q_1)\rangle}\times\frac{\langle \delta(q(0)-q_1)\rangle}{\langle \theta(q_1-q)\rangle}$$

$$k_{A\to B}(t)=\frac{\langle \dot{q}(0)\delta(q(0)-q_1)\theta(q(t)-q_1)\rangle}{\langle \delta(q(0)-q_1)\rangle}\times\frac{\langle \delta(q(0)-q_1)\rangle}{\langle \theta(q_1-q)\rangle}$$

Transmission coefficient

$$\kappa(t)\equiv\frac{k_{A\to B}(t)}{k_{A\to B}^{TST}}$$

$$=\frac{\langle \dot{q}(0)\delta(q(0)-q_1)\theta(q(t)-q_1)\rangle}{0.5\left|\dot{q}(0)\right|}$$

MD simulation to correct the transition state result!

MD simulation:

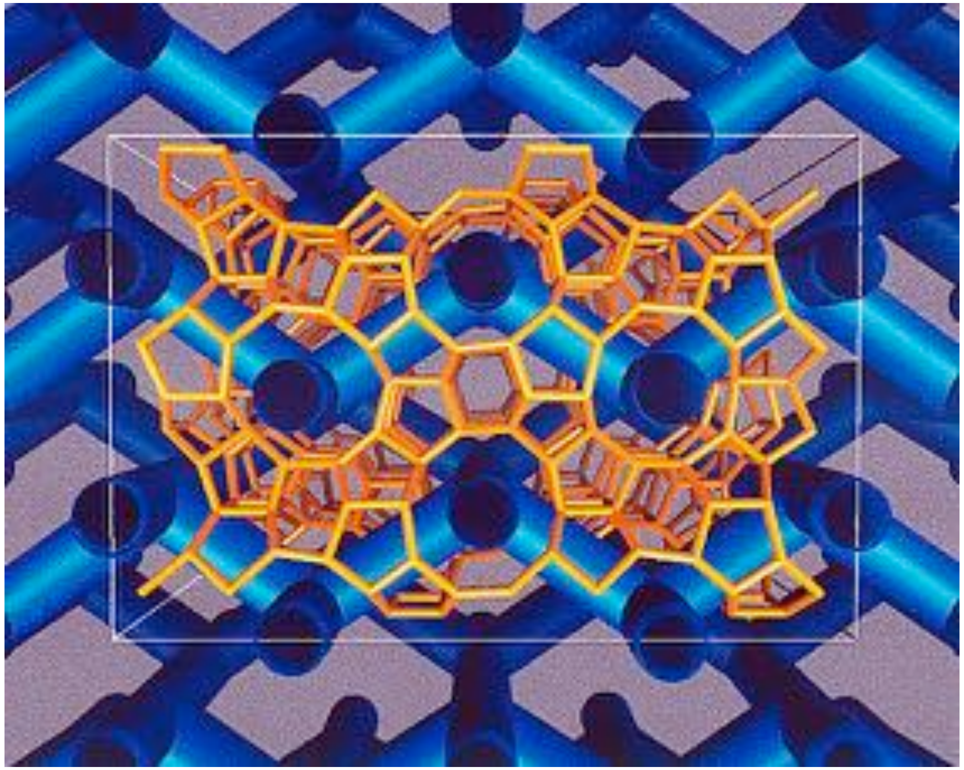
1. At t=0 $q=q_1$
2. Determine fraction at product state weighted with initial velocity

Example diffusion in zeolite

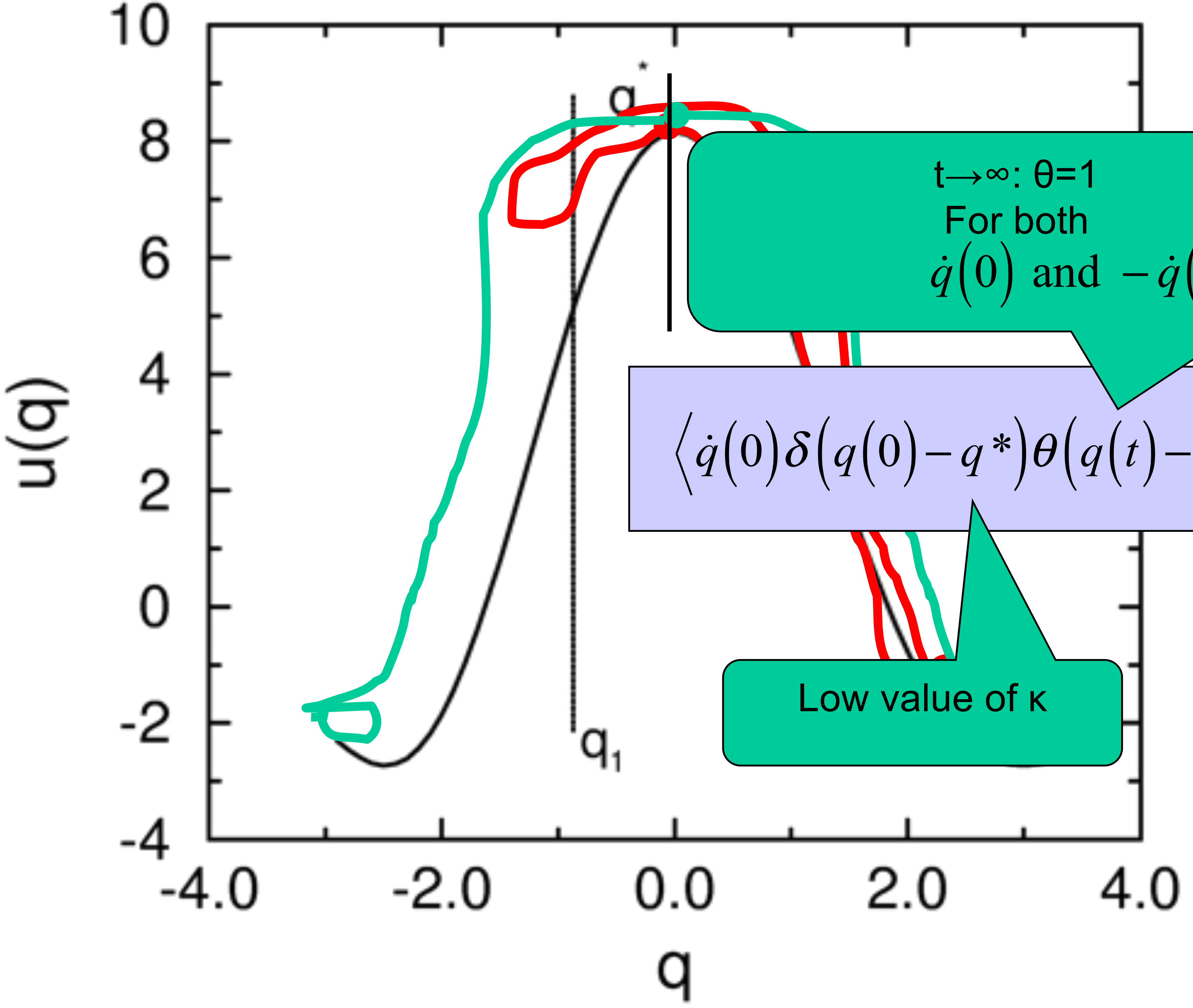
- Zeolites important class of materials

- Diffusion of alkanes in matrix is poorly described

- Approach
  - molecular simulation of alkanes in fixed zeolite frame
  - Unified atom FF by Dubbeldam et al.



D. Dubbeldam, et al., J. Phys. Chem. B, 108, 12301, 2004

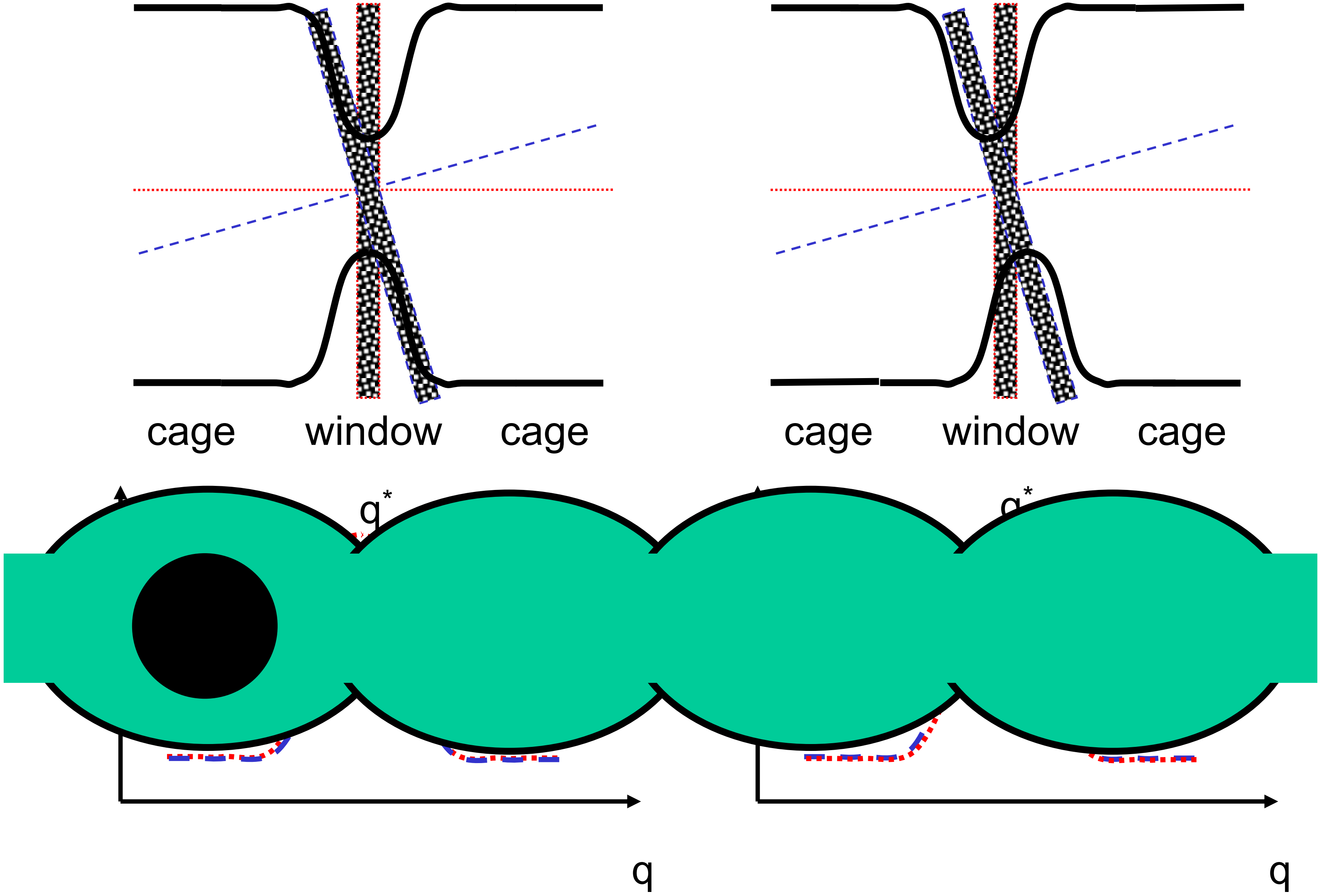


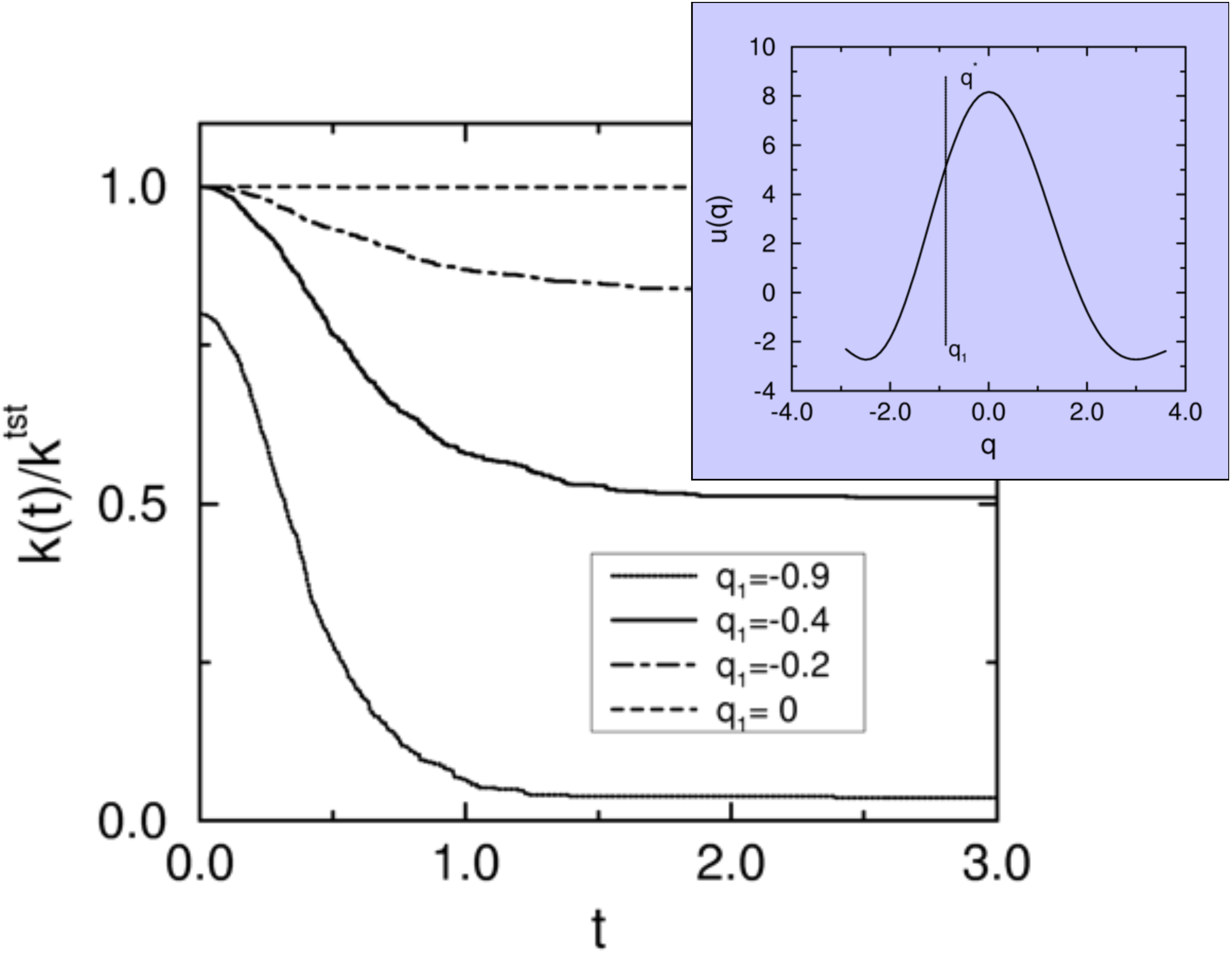
t→∞: θ=1
For both
$\dot{q}(0)$ and $-\dot{q}(0)$

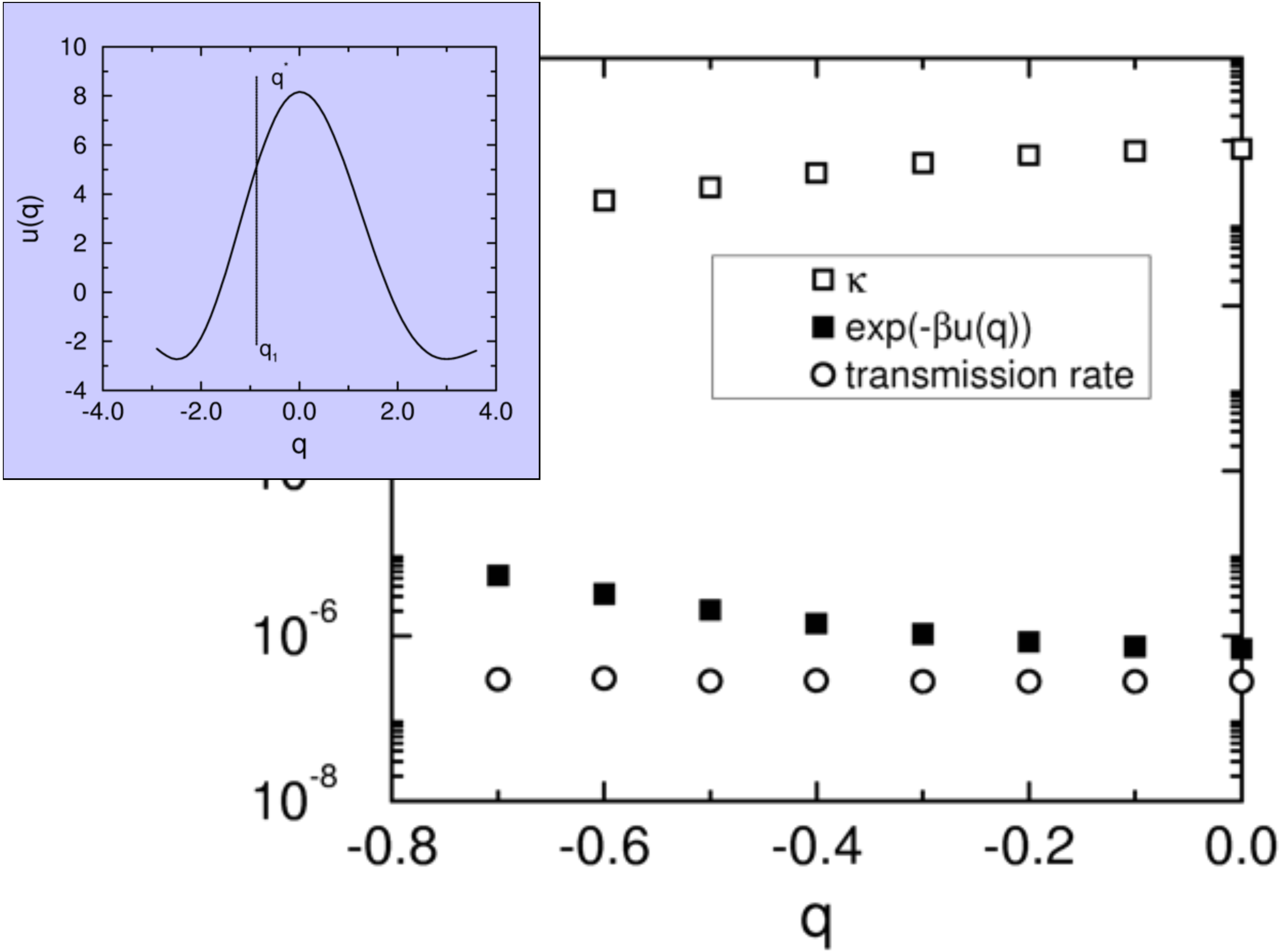
$$\left\langle \dot{q}(0)\delta\left(q(0)-q^*\right)\theta\left(q(t)-q^*\right)\right\rangle$$

Low value of κ

Reaction coordinate




k(t)/k^tst
1.0
0.5
0.0
0.0
1.0
2.0
3.0
t
q1=-0.9
q1=-0.4
q1=-0.2
q1= 0
u(q)
10
8
6
4
2
0
-2
-4
-4.0
-2.0
0.0
2.0
4.0
q
q*
q1


u(q)
10
8
6
4
2
0
-2
-4
-4.0
-2.0
0.0
2.0
4.0
q
q^*
q_1
10^{-6}
10^{-8}
-0.8
-0.6
-0.4
-0.2
0.0
q
□ κ
■ exp(-βu(q))
○ transmission rate

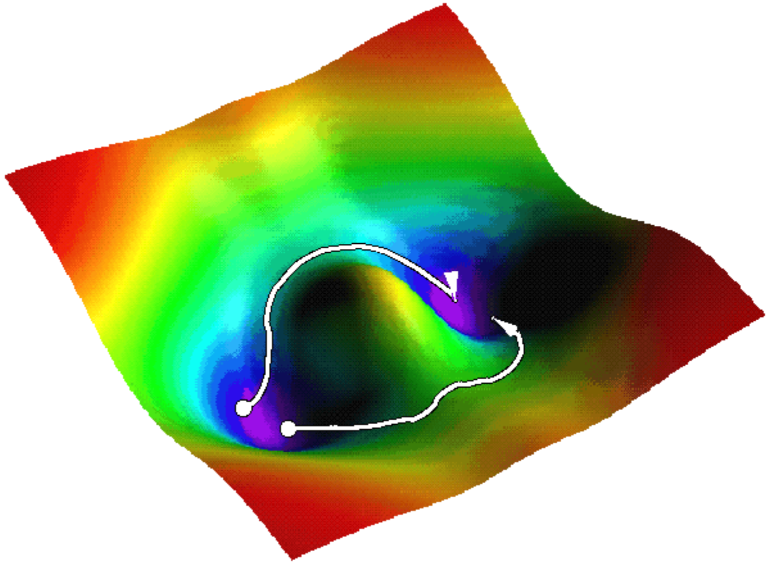
Outline

- Part 1
  - Rare event and reaction kinetics
  - Linear Response theory
  - Transition state theory
  - Free energy methods
  - Bennet Chandler approach
  - Example zeolites

- Part 2
  - **Two ended methods**
  - Transition path sampling
  - Rate constants
  - Reaction coordinate analysis
  - Application to crystallization
  - Path metadynamics

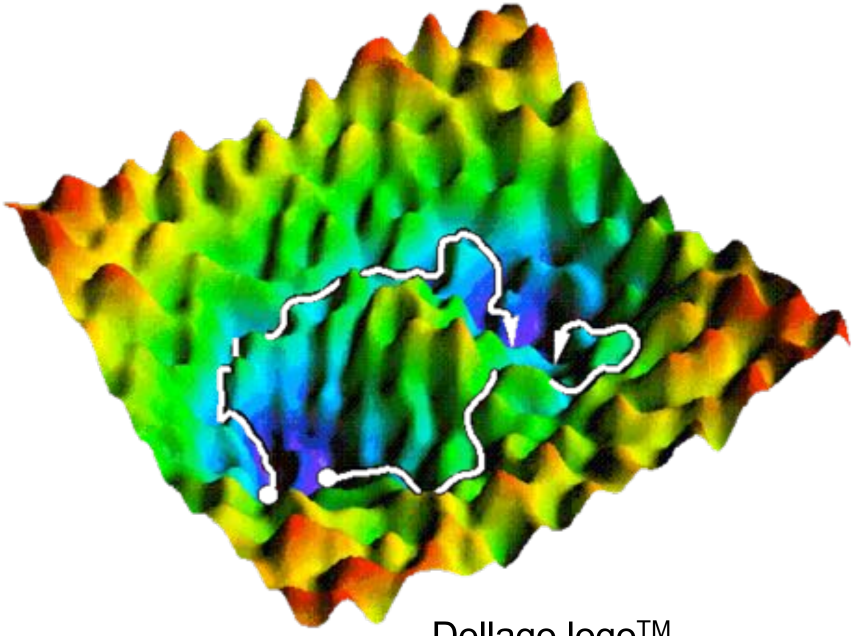
Barriers on smooth and rough energy landscapes

- Clearly, barrier is most important for rare event
- But how to obtain this barrier?
- In multidimensional energy landscapes barrier is saddle point



- # saddle points limited
- determined by potential energy
- use eigenvectors or Hessian to find them

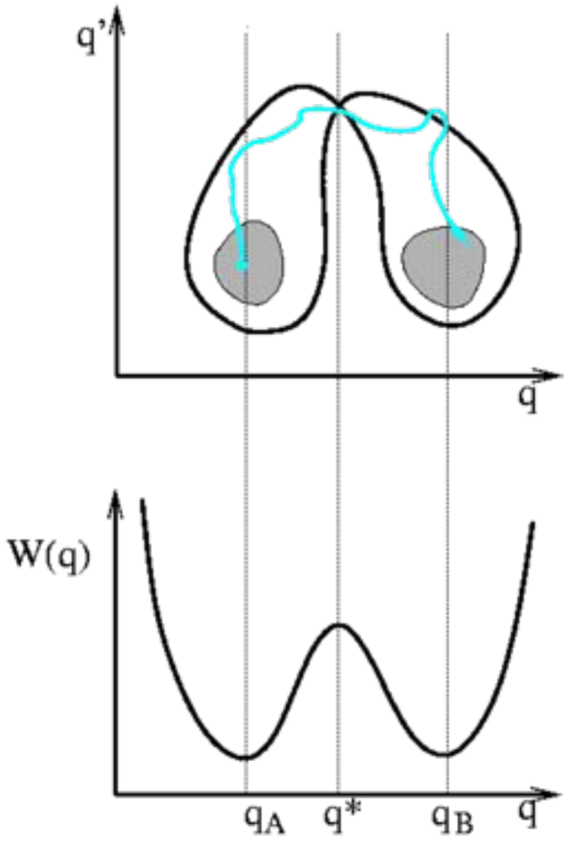
- # saddle points uncountable
- entropy important, many pathways
- determined by free energy
- exploring requires sampling schemes



Dellago logo™

Breakdown of BC approach

kappa can become immeasurable low if the reaction coordinate is at the wrong value the reaction coordinate is wrongly chosen

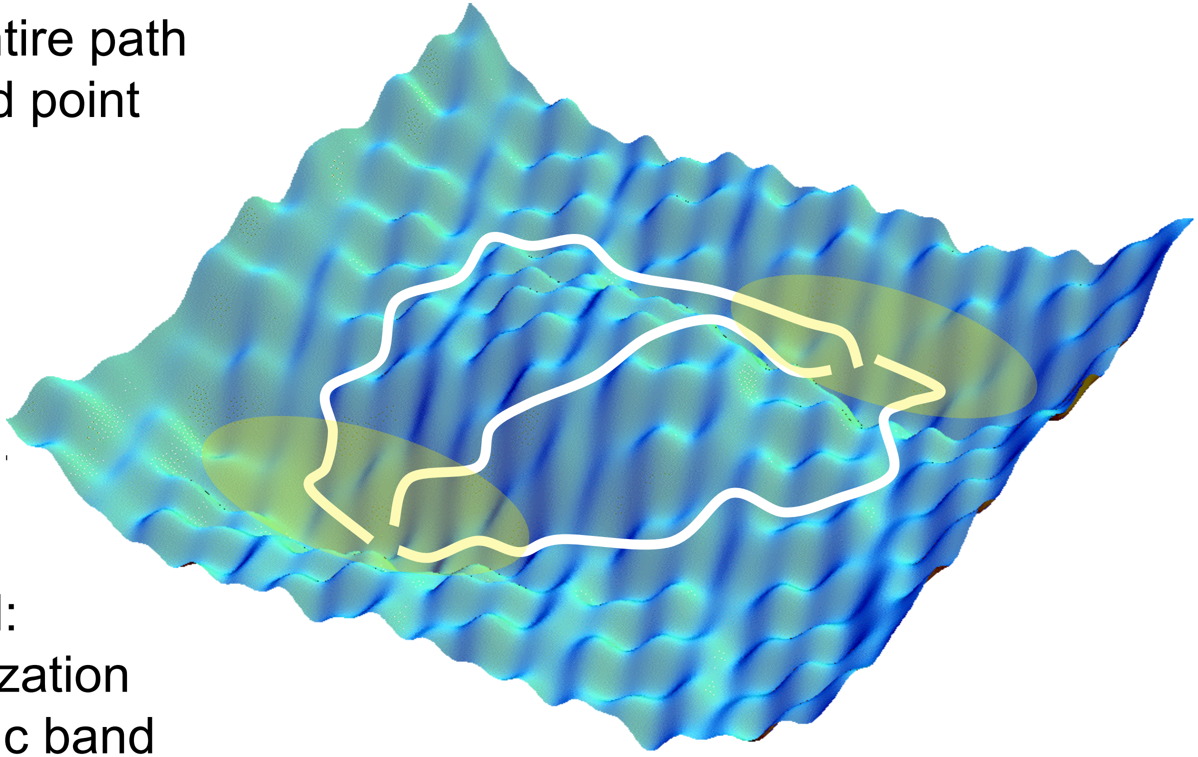


$$W(q) = -kT \ln \int dq' \exp\{-\beta E(q,q')\}$$

If the reaction coordinate is not known, the wrong order parameter can lead to wrong transition states, mechanism and rates

Two ended methods

Methods that take the entire path
and fix the begin and end point



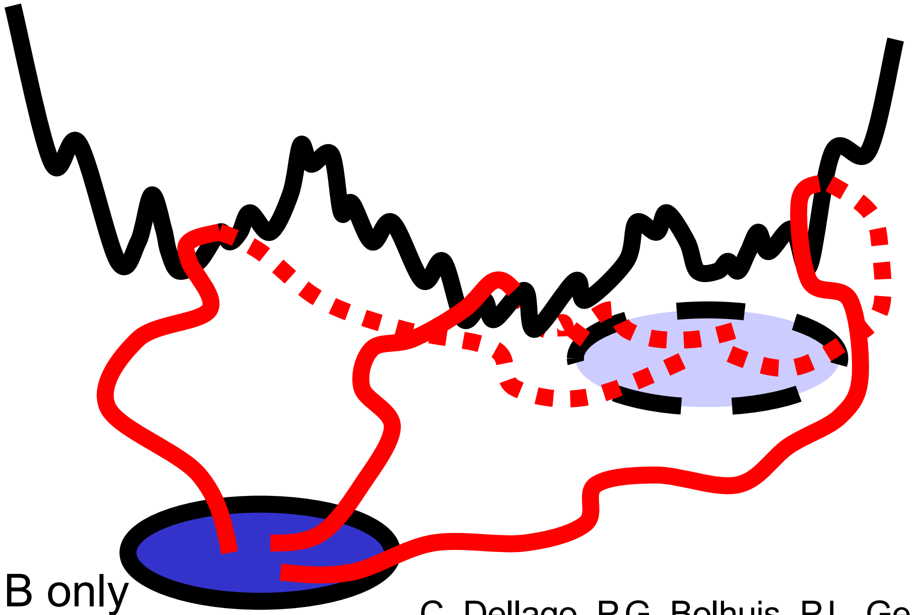
Many methods proposed:

Action minimization
Nudged elastic band
String method
Path metadynamics
Milestoning
Transition path sampling
....

Transition path sampling

Samples the path ensemble:

all trajectories that lead over barrier



- Sampling by Monte Carlo
- Requires definition of stable states A,B only
- Results in ensemble of pathways
- Reaction coordinate is a result of simulation not an input
- Allows for calculation of rate constants

C. Dellago, P.G. Bolhuis, P.L. Geissler

Adv. Chem. Phys. **123**, 1 2002

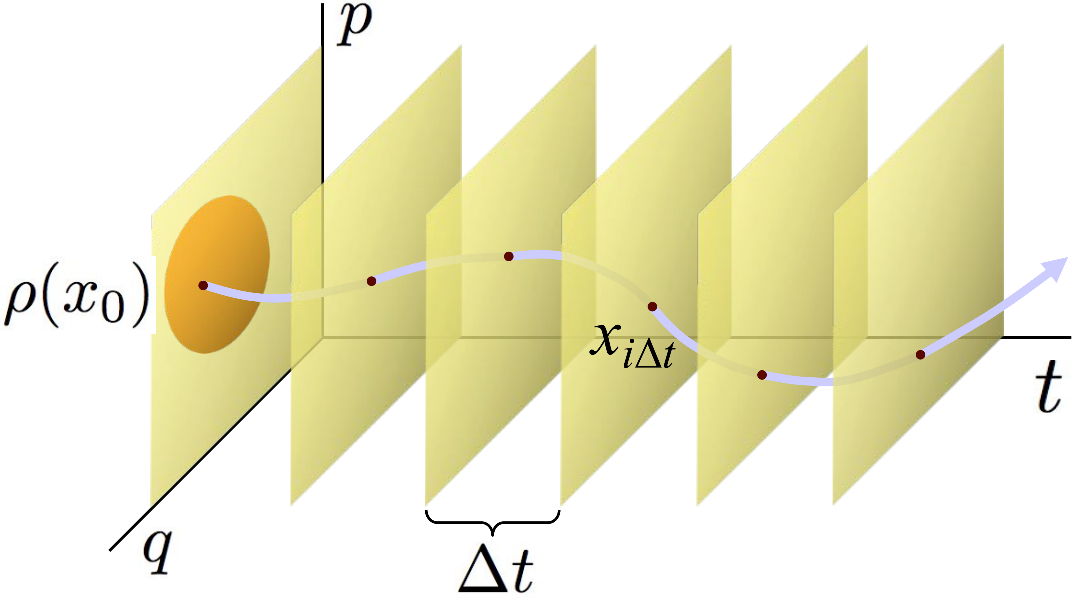
Apply when process of interest

- is a rare event
- is complex and reaction coordinate is not known

Examples: nucleation, reactions in solution, protein folding

Path probability density

$x(\mathcal{T}) \equiv \{x_0, x_{\Delta t}, x_{2\Delta t}, \ldots, x_{\mathcal{T}}\}$ Path = Sequence of states



$$\mathcal{P}[x(\mathcal{T})] = \rho(x_0) \prod_{i=0}^{\mathcal{T}/\Delta t - 1} p(x_{i\Delta t} \rightarrow x_{(i+1)\Delta t})$$

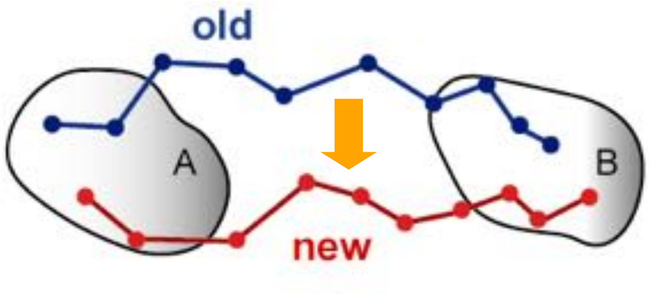
Transition path ensemble

$$\boxed{\mathcal{P}_{AB}[x(\mathcal{T})] \equiv h_A(x_0)\mathcal{P}[x(\mathcal{T})]h_B(x_{\mathcal{T}})/Z_{AB}(\mathcal{T})}$$

$$Z_{AB}(\mathcal{T}) \equiv \int \mathcal{D}x(\mathcal{T})\, h_A(x_0)\mathcal{P}[x(\mathcal{T})]h_B(x_{\mathcal{T}})$$

$$\int \mathcal{D}x(\mathcal{T}) \equiv \int \cdots \int dx_0 dx_{\Delta t} dx_{2\Delta t} \cdots dx_{\mathcal{T}}$$

Metropolis MC of pathways



1. Generate **new** path from **old** one

$$x^{(o)}(\mathcal{T}) \longrightarrow x^{(n)}(\mathcal{T})$$

2. Accept **new** path according to **detailed balance**:

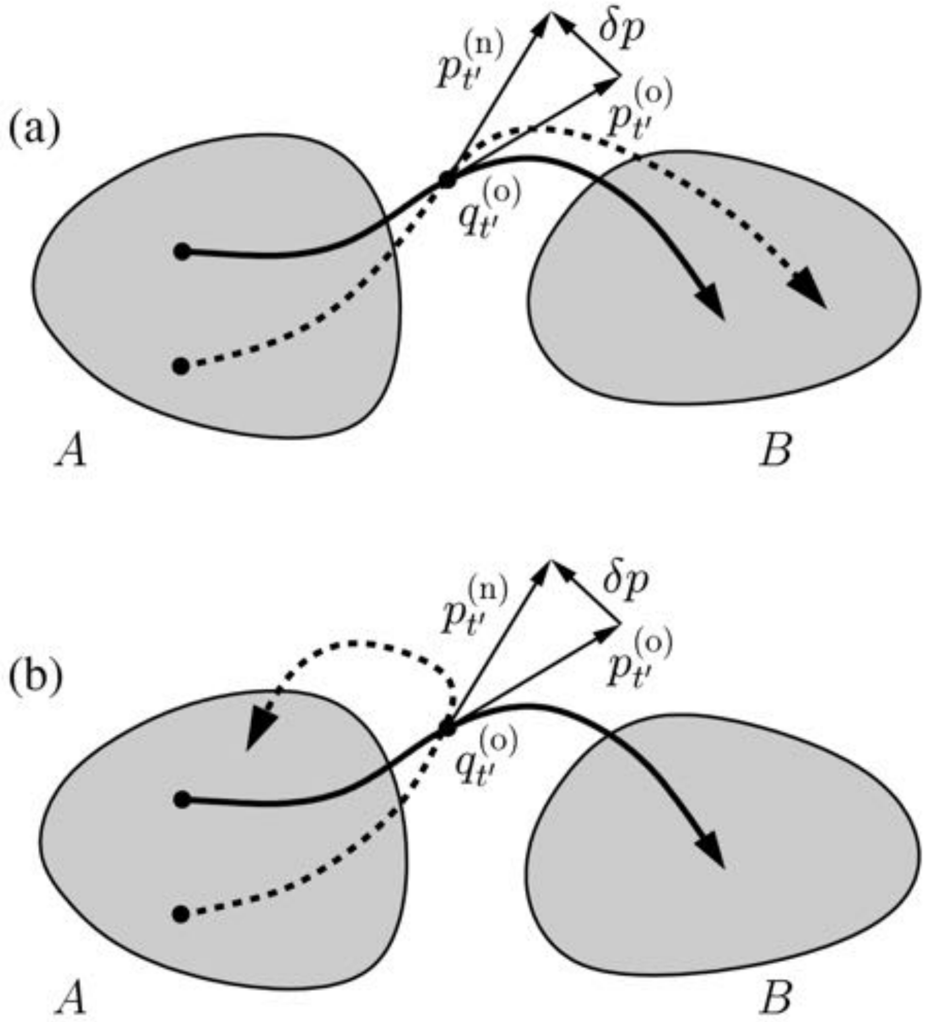
$$\mathcal{P}_{AB}[x^{(\mathrm{o})}(\mathcal{T})]\pi[x^{(\mathrm{o})}(\mathcal{T}) \to x^{(\mathrm{n})}(\mathcal{T})] = \mathcal{P}_{AB}[x^{(\mathrm{n})}(\mathcal{T})]\pi[x^{(\mathrm{n})}(\mathcal{T}) \to x^{(\mathrm{o})}(\mathcal{T})]$$

$$\pi[x^{(\mathrm{o})}(\mathcal{T}) \to x^{(\mathrm{n})}(\mathcal{T})] = P_{\mathrm{gen}}[x^{(\mathrm{o})}(\mathcal{T}) \to x^{(\mathrm{n})}(\mathcal{T})] \times P_{\mathrm{acc}}[x^{(\mathrm{o})}(\mathcal{T}) \to x^{(\mathrm{n})}(\mathcal{T})]$$

3. Satisfy detailed balance with the **Metropolis rule**:

$$P_{\mathrm{acc}}[x^{(\mathrm{o})}(\mathcal{T}) \to x^{(\mathrm{n})}(\mathcal{T})] = h_A[x_0^{(\mathrm{n})}]h_B[x_{\mathcal{T}}^{(\mathrm{n})}]\min\left\{1, \frac{\mathcal{P}[x^{(\mathrm{n})}(\mathcal{T})]P_{\mathrm{gen}}[x^{(\mathrm{n})}(\mathcal{T}) \to x^{(\mathrm{o})}(\mathcal{T})]}{\mathcal{P}[x^{(\mathrm{o})}(\mathcal{T})]P_{\mathrm{gen}}[x^{(\mathrm{o})}(\mathcal{T}) \to x^{(\mathrm{n})}(\mathcal{T})]}\right\}$$

Shooting moves



accept

reject

$$P_{acc}[x^{(o)}(T) \to x^{(n)}(T)] = h_A(x_0^{(n)})h_B(x_T^{(n)})$$

$$h_A(t) = \begin{cases} 1 & if \quad x_t \in A \\ 0 & if \quad x_t \notin A \end{cases}$$

Shooting algorithm

$$P^{\mathrm{f}}_{\mathrm{gen}}[x^{\mathrm{o}}(\mathcal{T}) \rightarrow x^{\mathrm{n}}(\mathcal{T})] = \prod_{i=t'/\Delta t}^{\mathcal{T}/\Delta t-1} p\left(x^{(\mathrm{n})}_{i\Delta t} \rightarrow x^{(\mathrm{n})}_{(i+1)\Delta t}\right)$$

$$P^{\mathrm{b}}_{\mathrm{gen}}[x^{\mathrm{o}}(\mathcal{T}) \rightarrow x^{\mathrm{n}}(\mathcal{T})] = \prod_{i=1}^{t'/\Delta t} \bar{p}\left(x^{(\mathrm{n})}_{i\Delta t} \rightarrow x^{(\mathrm{n})}_{(i-1)\Delta t}\right)$$

$$P_{\mathrm{gen}}[x^{(\mathrm{o})}(\mathcal{T}) \rightarrow x^{(\mathrm{n})}(\mathcal{T})] = p_{\mathrm{gen}}[x^{(\mathrm{o})}_{t'} \rightarrow x^{(\mathrm{n})}_{t'}] \prod_{i=t'/\Delta t}^{\mathcal{T}/\Delta t-1} p\left(x^{(\mathrm{n})}_{i\Delta t} \rightarrow x^{(\mathrm{n})}_{(i+1)\Delta t}\right) \times \prod_{i=1}^{t'/\Delta t} \bar{p}\left(x^{(\mathrm{n})}_{i\Delta t} \rightarrow x^{(\mathrm{n})}_{(i-1)\Delta t}\right)$$

$$P_{\mathrm{acc}}[x^{(\mathrm{o})}(\mathcal{T}) \rightarrow x^{(\mathrm{n})}(\mathcal{T})] = h_A[x^{(\mathrm{n})}_0] h_B[x^{(\mathrm{n})}_{\mathcal{T}}] \min\left[1, \frac{\rho\left(x^{(\mathrm{n})}_0\right)}{\rho\left(x^{(\mathrm{o})}_0\right)} \prod_{i=0}^{t'/\Delta t-1} \frac{p\left(x^{(\mathrm{n})}_{i\Delta t} \rightarrow x^{(\mathrm{n})}_{(i+1)\Delta t}\right)}{\bar{p}\left(x^{(\mathrm{n})}_{(i+1)\Delta t} \rightarrow x^{(\mathrm{n})}_{i\Delta t}\right)} \times \frac{\bar{p}\left(x^{(\mathrm{o})}_{(i+1)\Delta t} \rightarrow x^{(\mathrm{o})}_{i\Delta t}\right)}{p\left(x^{(\mathrm{o})}_{i\Delta t} \rightarrow x^{(\mathrm{o})}_{(i+1)\Delta t}\right)}\right]$$

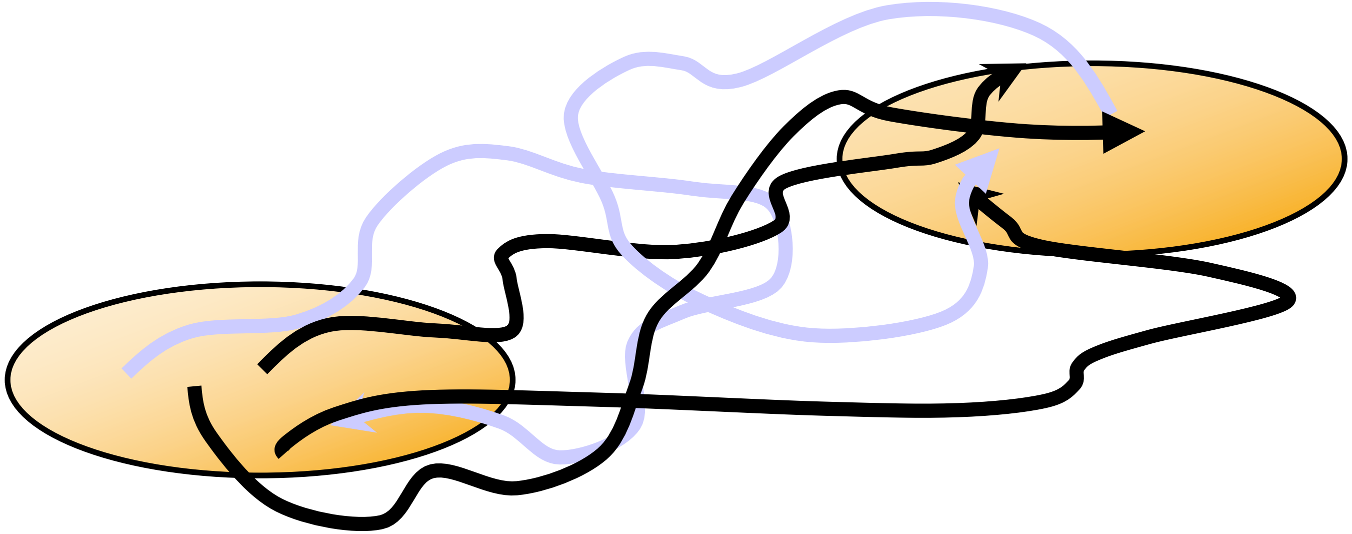
$$\frac{p(x \rightarrow y)}{\bar{p}(y \rightarrow x)} = \frac{\rho_0(y)}{\rho_0(x)}$$

$$P_{\mathrm{acc}}[x^{(\mathrm{o})}(\mathcal{T}) \rightarrow x^{(\mathrm{n})}(\mathcal{T})] = h_A[x^{(\mathrm{n})}_0] h_B[x^{(\mathrm{n})}_{\mathcal{T}}] \min\left[1, \frac{\rho(x^{(\mathrm{n})}_{t'})}{\rho(x^{(\mathrm{o})}_{t'})}\right]$$

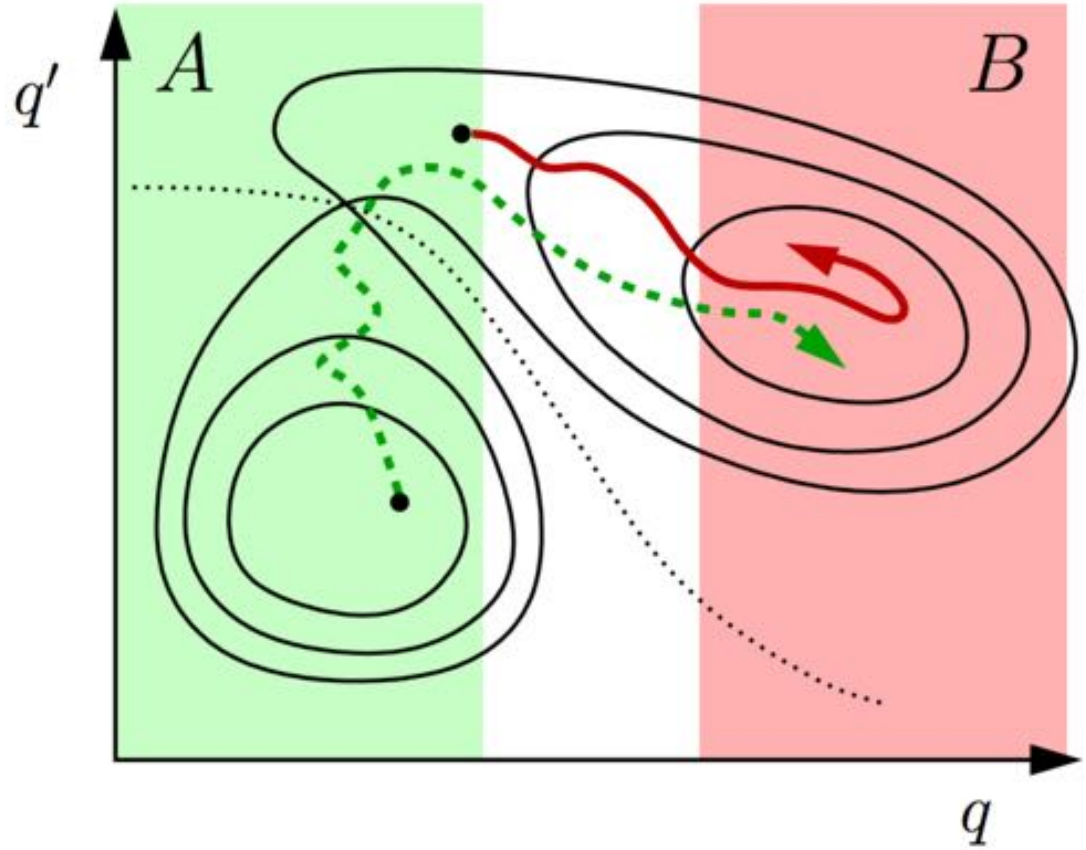
$$P_{\mathrm{acc}}[x^{(\mathrm{o})}(\mathcal{T}) \rightarrow x^{(\mathrm{n})}(\mathcal{T})] = h_A[x^{(\mathrm{n})}_0] h_B[x^{(\mathrm{n})}_{\mathcal{T}}]$$

Standard TPS algorithm

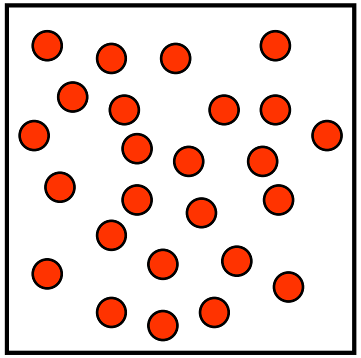
- take existing path
- choose random time slice t
- change momenta slightly at t
- integrate forward and backward in time to create new path of length L
- accept if A and B are connected, otherwise reject and retain old path
- calculate averages
- repeat



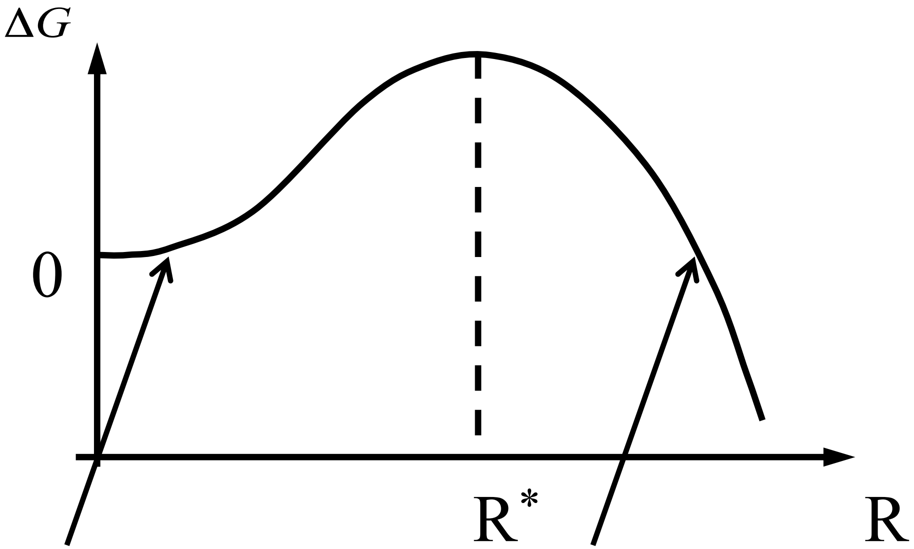
Definition of the stable states

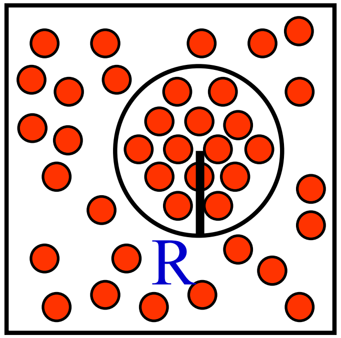


Classical nucleation (1926)



Liquid





Crystal nucleus

surface bulk

$$\Delta G = 4\pi R^2 \gamma - \frac{4}{3}\pi R^3 \rho \Delta\mu_{ls}$$

–How does the crystal form?

–What is the structure of the critical nucleus

–Is classical nucleation theory correct?

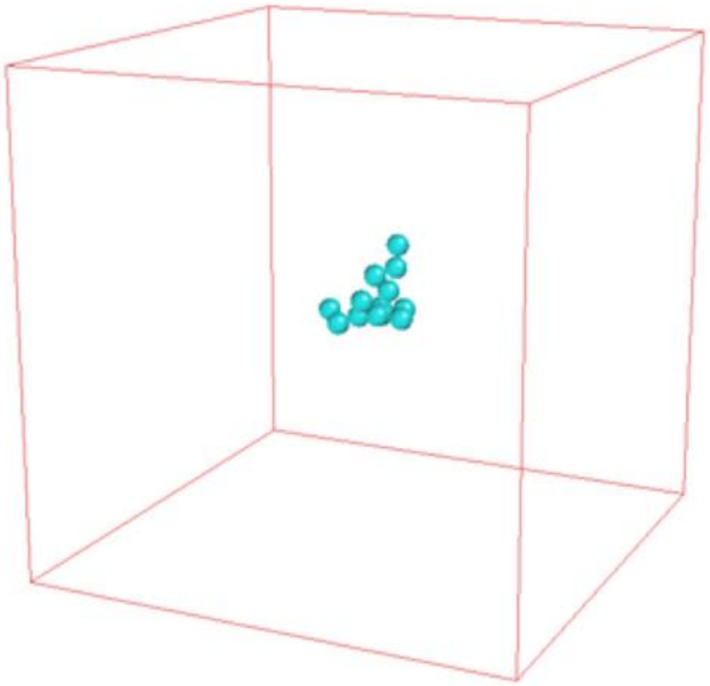
- What is the barrier?
- Rate constant

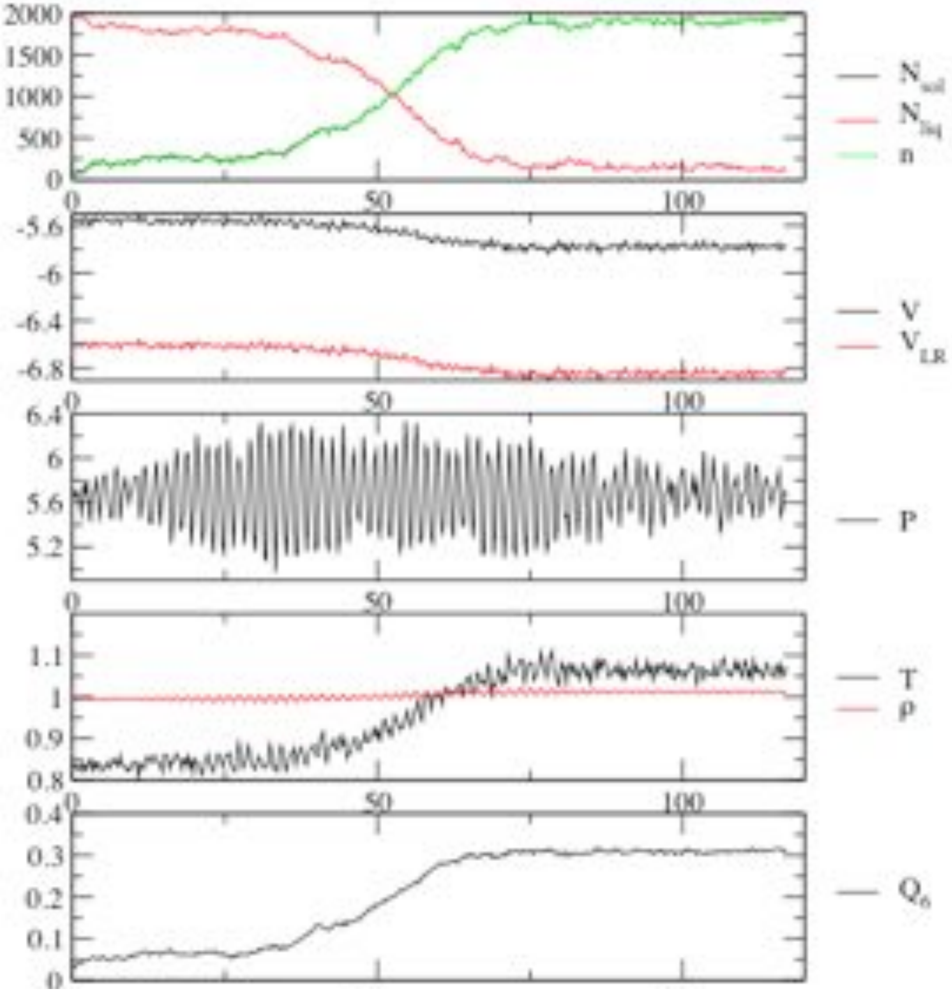
γ : surface tension
$\Delta\mu$: chem. pot difference
ρ: density

Path sampling of nucleation

TIS in NPH ensemble, as density and temperature change

N=10000, P=5.68 H=1.41 (25 % undercooling)





D. Moroni, P. R. ten Wolde, and P. G. Bolhuis, *Phys. Rev. Lett.* **94**, 235703 (2005)

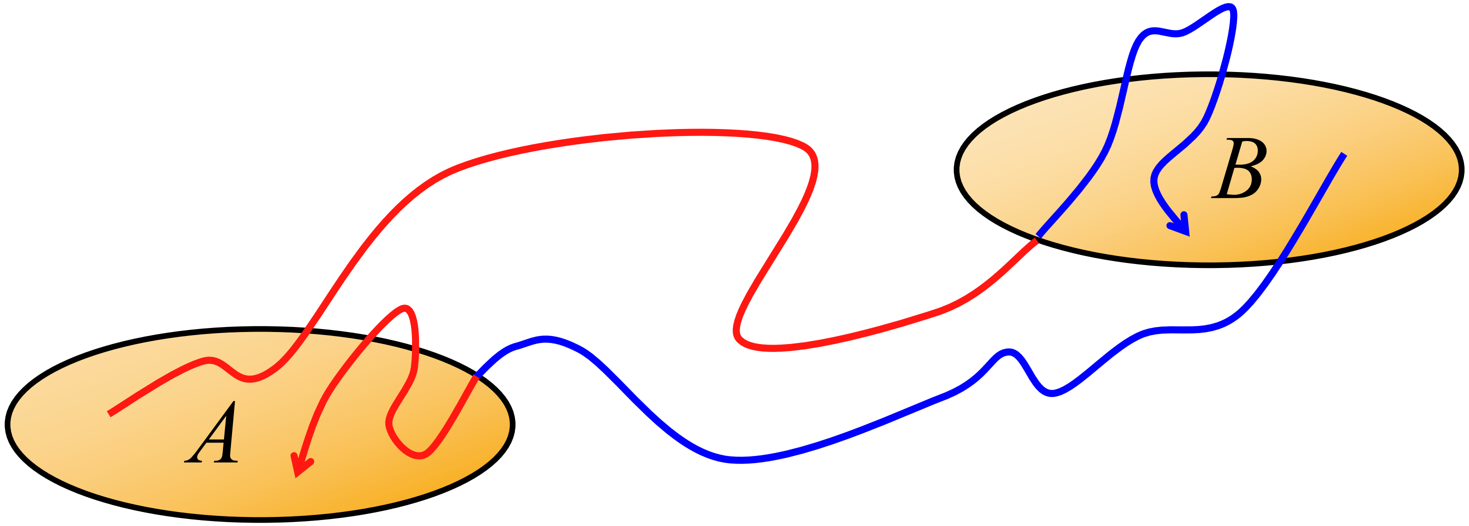
Sampling paths is only the beginning

- *Eugene Wigner:* ***"It is nice to know that the computer understands the problem. But I would like to understand it too."***

- Path ensemble needs to be further explored to obtain:
  - Rate constants
  - Free energy
  - Transition state ensembles
  - Mechanistic picture
  - Reaction coordinate

- Illustrative example: crystal nucleation

Transition interface sampling



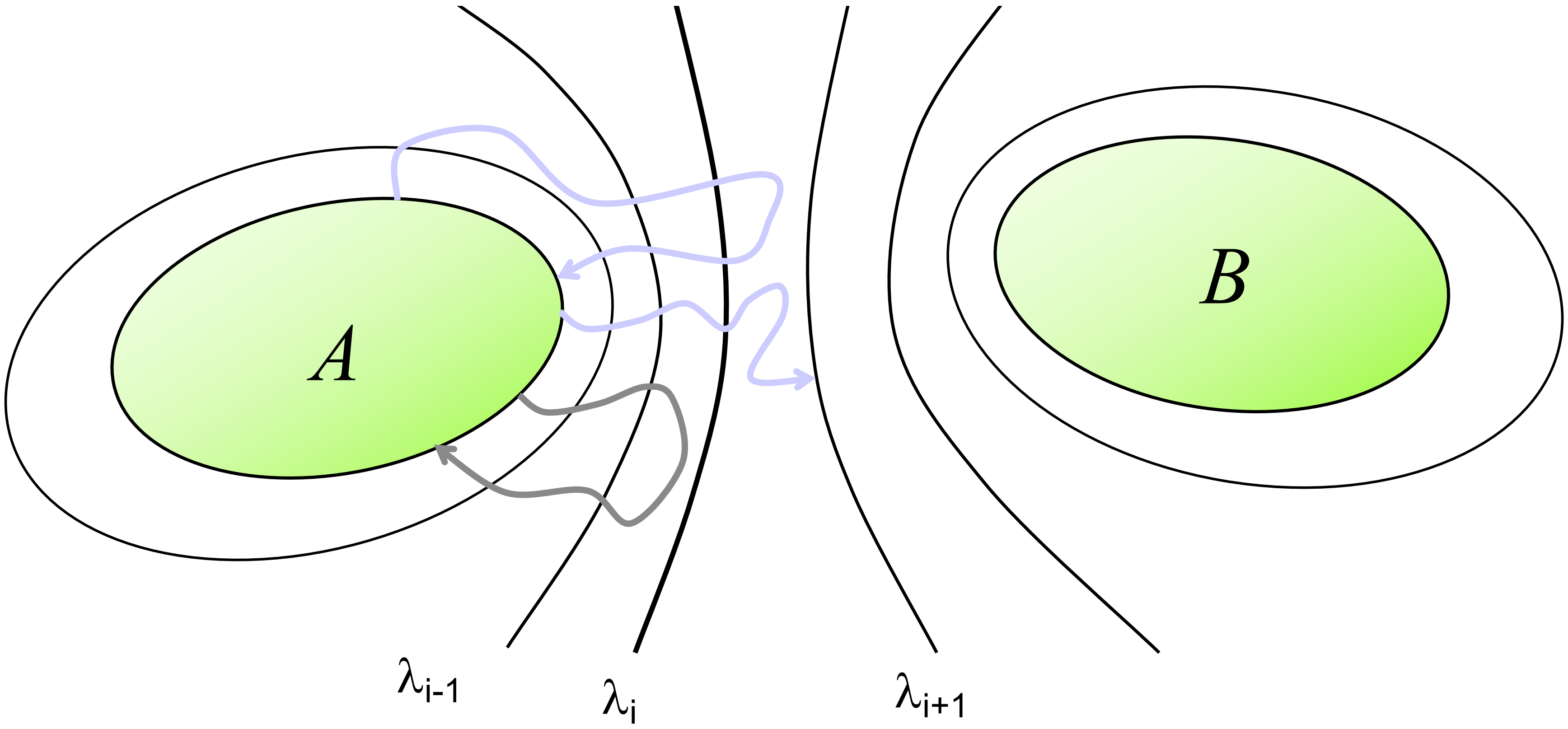
Overall states in phase space: $\mathcal{A}$ directly coming from **A**

$\mathcal{B}$ directly coming from **B**

$$C(t) \equiv \frac{\langle h_{\mathcal{A}}(x_0) h_{\mathcal{B}}(x_t) \rangle}{\langle h_{\mathcal{A}} \rangle}$$

$$k_{AB} = \frac{\langle h_{\mathcal{A}}(x_0) \dot{h}_{\mathcal{B}}(x_0) \rangle}{\langle h_{\mathcal{A}} \rangle} = \frac{\langle \phi_{AB} \rangle}{\langle h_{\mathcal{A}} \rangle}$$

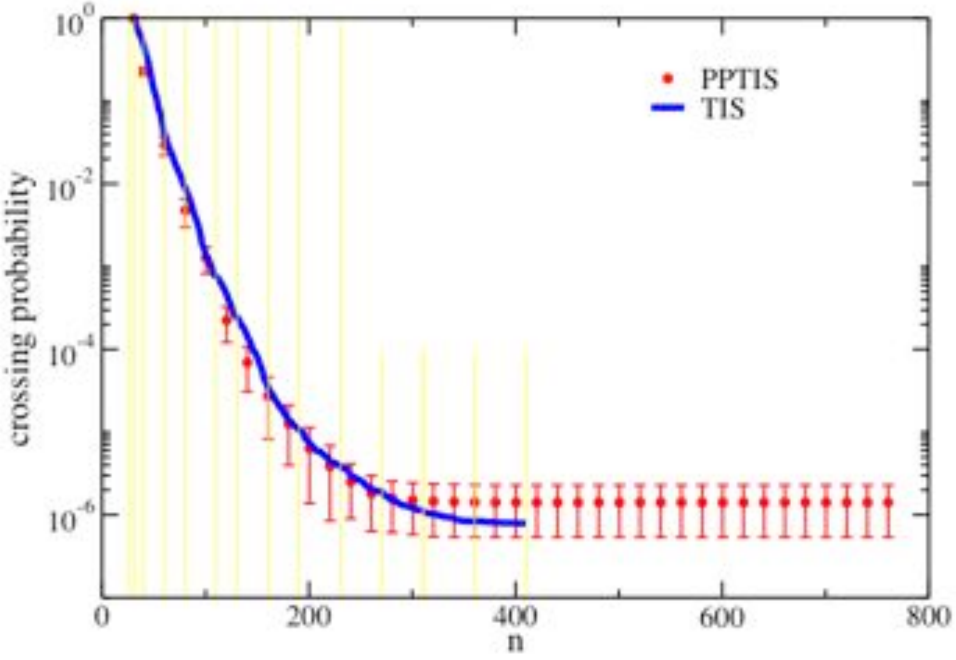
T. S. van Erp, D. Moroni and P. G. Bolhuis, *J. Chem. Phys*. **118** , 7762 (2003)
T. S. van Erp and P. G. Bolhuis, *J. Comp. Phys.* **205**, 157 (2005)



$P_A(\lambda_{i+1} | \lambda_i)$= probability that path crossing i for first time after leaving A reaches i+1 before A

$$k_{AB}^{TIS} = \frac{\langle \phi_{AB} \rangle}{\langle h_{\mathcal{A}} \rangle} = \frac{\langle \phi_{AB} \rangle}{\langle h_{\mathcal{A}} \rangle} \prod_{i=1}^{n-1} P_A(\lambda_{i+1} | \lambda_i) = \Phi_A \prod_{i=1}^{n-1} P_A(\lambda_{i+1} | \lambda_i)$$

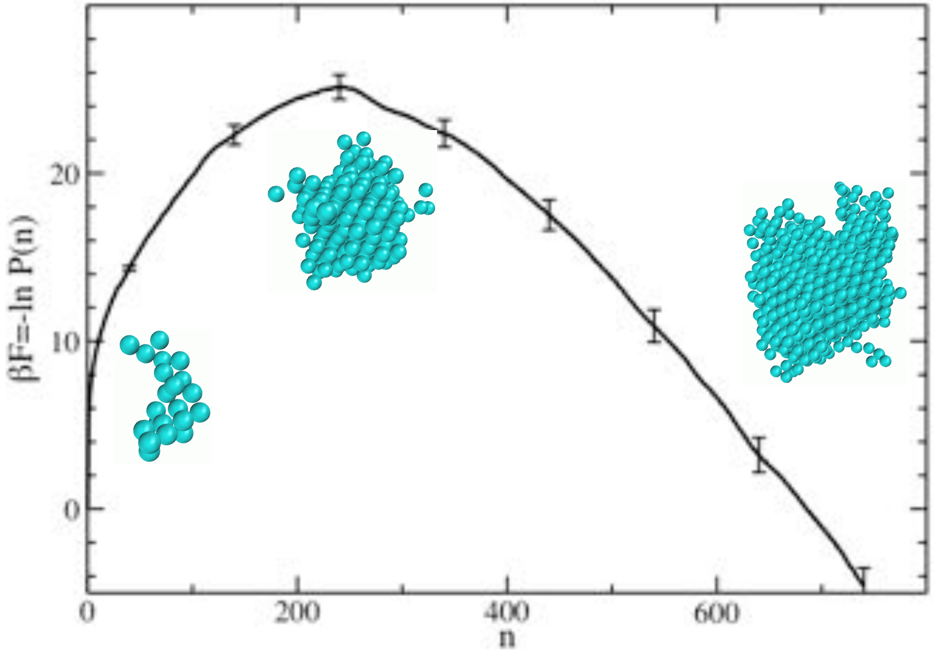
TIS results for nucleation



$$\mathcal{P}_A(B|1) = \prod_{i=1}^{n-1} \mathcal{P}_A(i+1|i) = 8 \times 10^{-7}$$

$$\frac{\langle \phi_{A1} \rangle}{\langle h_{\mathcal{A}} \rangle} = 1.29$$

$$k_{AB} = (1.0 \pm 0.8) \times 10^{-6}$$



Free energy follows directly

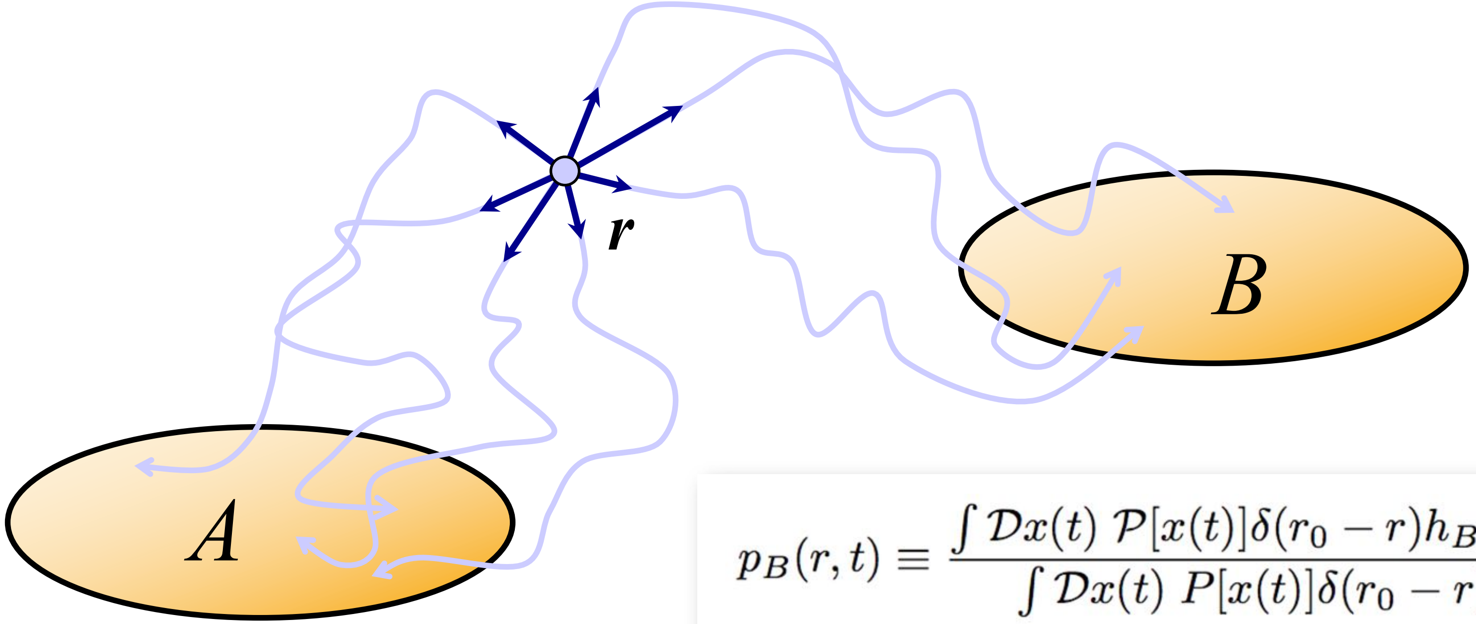
Moroni, van Erp, Bolhuis, PRE, 2005

Structural analysis?

Committor

(aka p-fold, splitting probability)

$p_B(r,t) =$ Probability that a trajectory initiated at r relaxes into B



$$p_B(r,t) \equiv \frac{\int \mathcal{D}x(t)\ \mathcal{P}[x(t)]\delta(r_0 - r)h_B(x_t)}{\int \mathcal{D}x(t)\ P[x(t)]\delta(r_0 - r)}$$

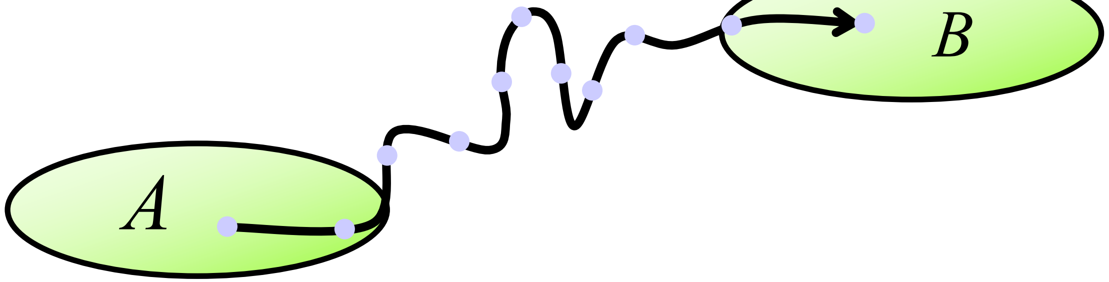
$$p_B(r,t) \approx \frac{1}{N}\sum_{i=1}^{N} h_B(x_t^{(i)})$$

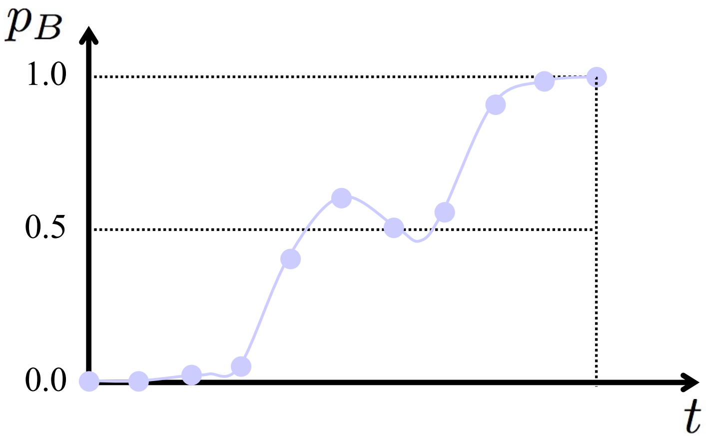
$$\sigma = \sqrt{\langle (p_B^{(N)} - p_B)^2 \rangle} = \sqrt{\frac{p_B(1-p_B)}{N}}$$

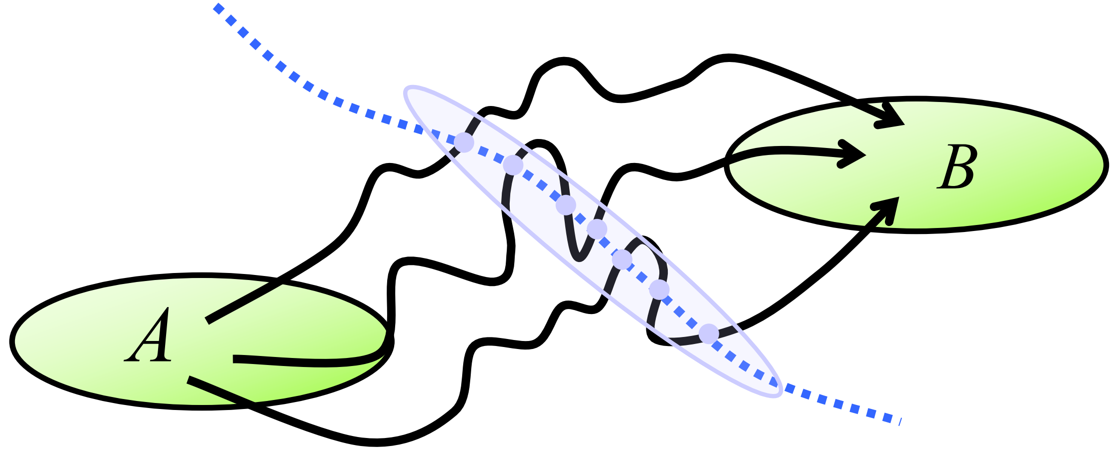
L. Onsager, *Phys. Rev.* **54**, 554 (1938).
M. M. Klosek, B. J. Matkowsky, Z. Schuss, *Ber. Bunsenges. Phys. Chem.* **95**, 331 (1991)
V. Pande, A. Y. Grosberg, T. Tanaka, E. I. Shaknovich, *J. Chem. Phys.* **108**, 334 (1998)

Transition state ensemble

r is a **transition state** (TS) if $p_B(r) = p_A(r) = 0.5$



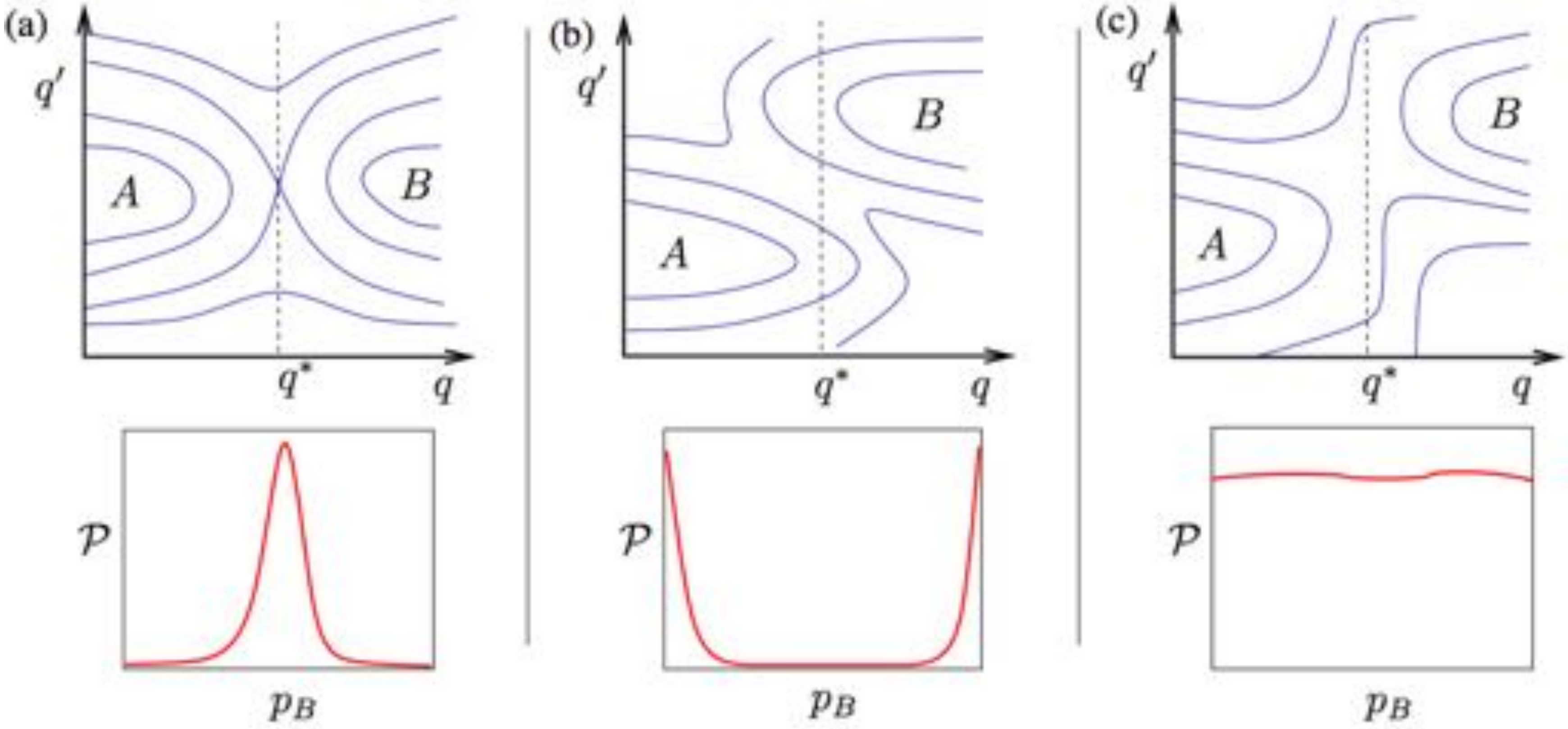




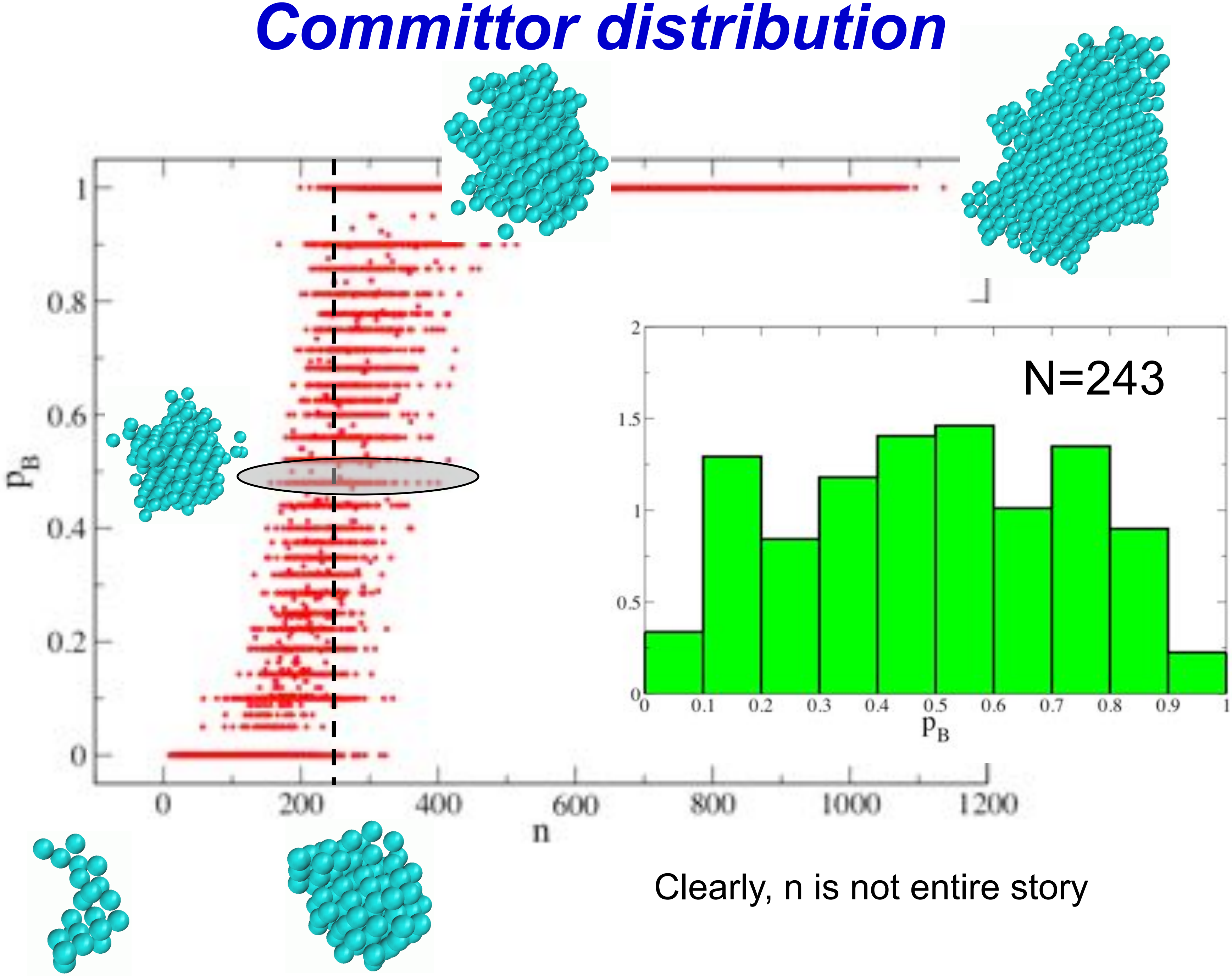
TSE:

Intersections of transition pathways with the p_B=1/2 surface

Committor distributions



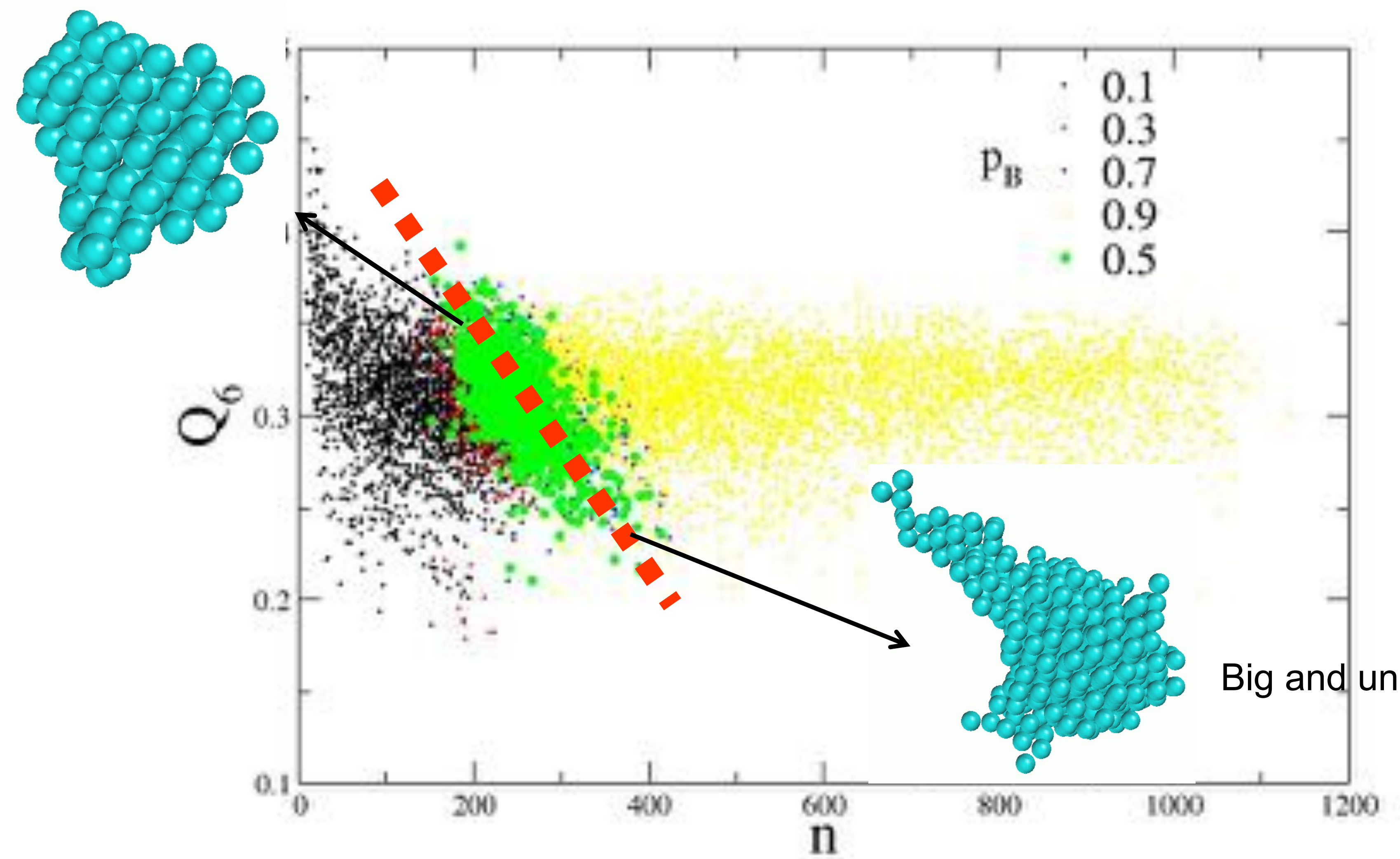
Committor distribution



Clearly, n is not entire story

Structure

Small and structured



Big and unstructured

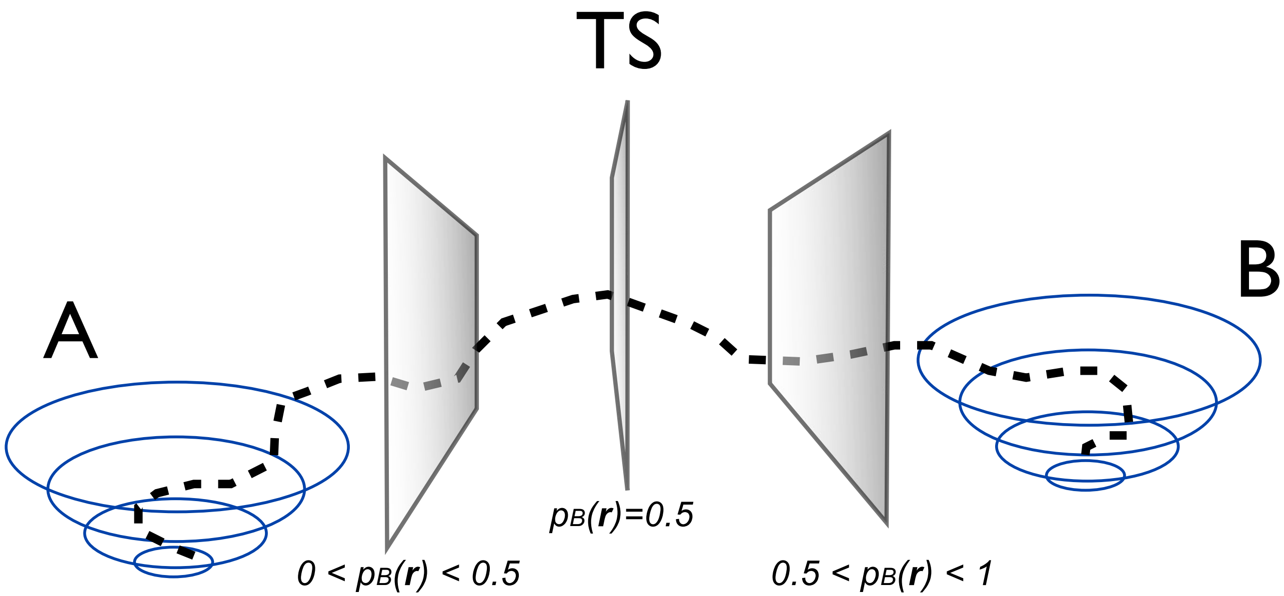
Committor analysis gives valuable insight

Outline

- Part 1
  - Rare event and reaction kinetics
  - Linear Response theory
  - Transition state theory
  - Free energy methods
  - Bennet Chandler approach
  - Example zeolites

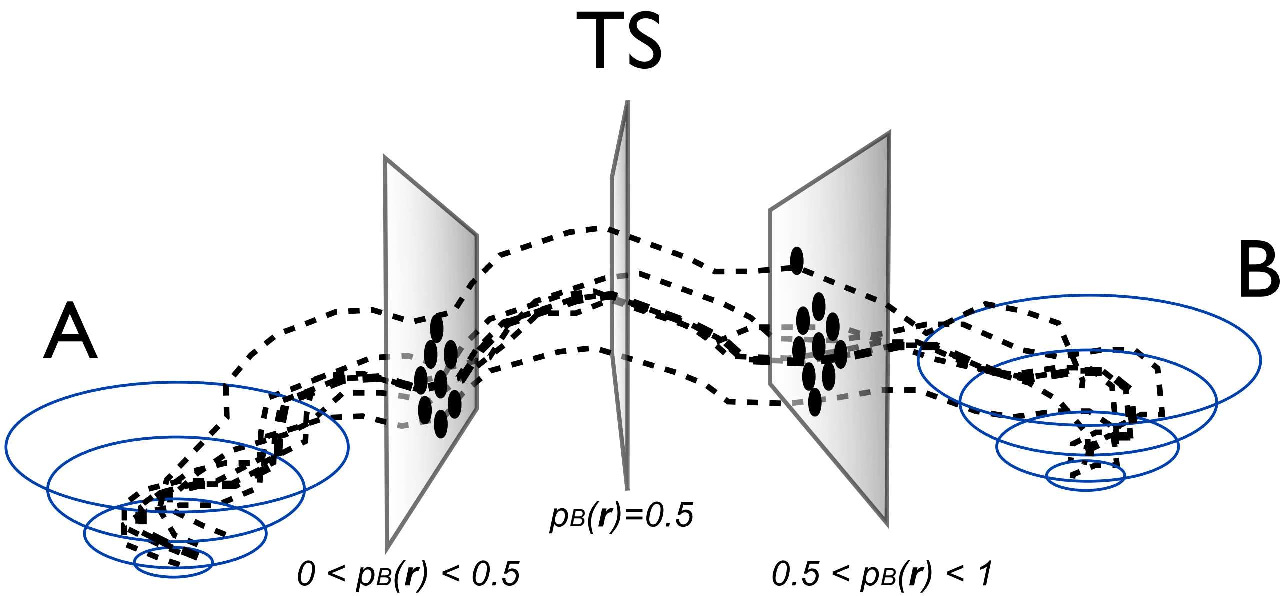
- Part 2
  - Two ended methods
  - Transition path sampling
  - Rate constants
  - Reaction coordinate analysis
  - Application to crystallization
  - **Path metadynamics**

Path-metadynamics



iso-committor surfaces

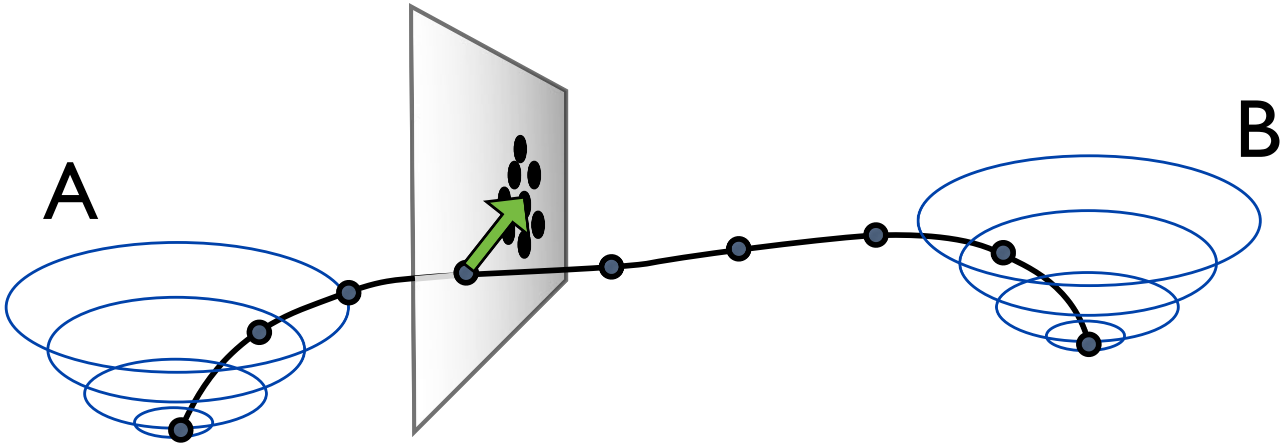
Path-metadynamics



the average transition pathway
(in CV space)

Path finding on high-dimensional free energy landscapes.
Grisell Díaz Leines and Bernd Ensing
Phys. Rev. Lett. ***109*** *(2012), 020601*

Path-metadynamics



distance to mean density

- start from guess path
- bias dynamics along path
- move nodes to the mean density
- maintain equidistant nodes

Path finding on high-dimensional free energy landscapes.
Grisell Díaz Leines and Bernd Ensing
Phys. Rev. Lett. ***109*** *(2012), 020601*

Path-metadynamics of alanine dipeptide

Parameters:

T = 300 K

$H_{gaussian}$ = 10 K

$W_{gaussian}$ = 0.05

$\Delta t_{gaussian}$ = 100 MD steps

n = 20 + 20 + 20

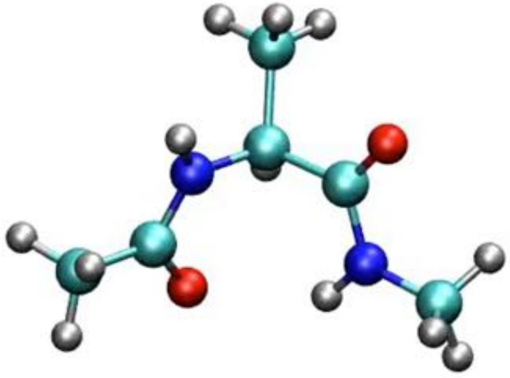
τ = 1000 MD steps

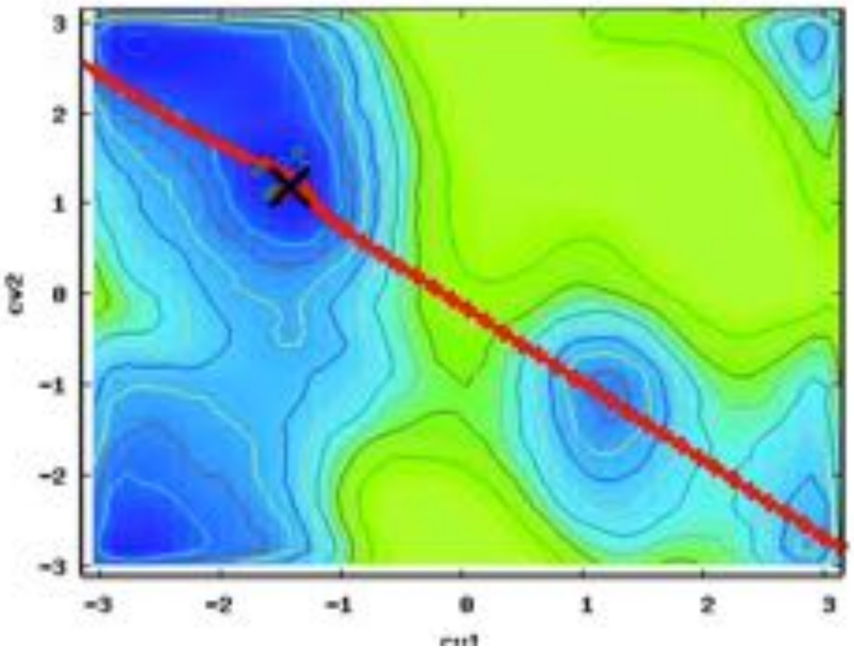
every recrossing:

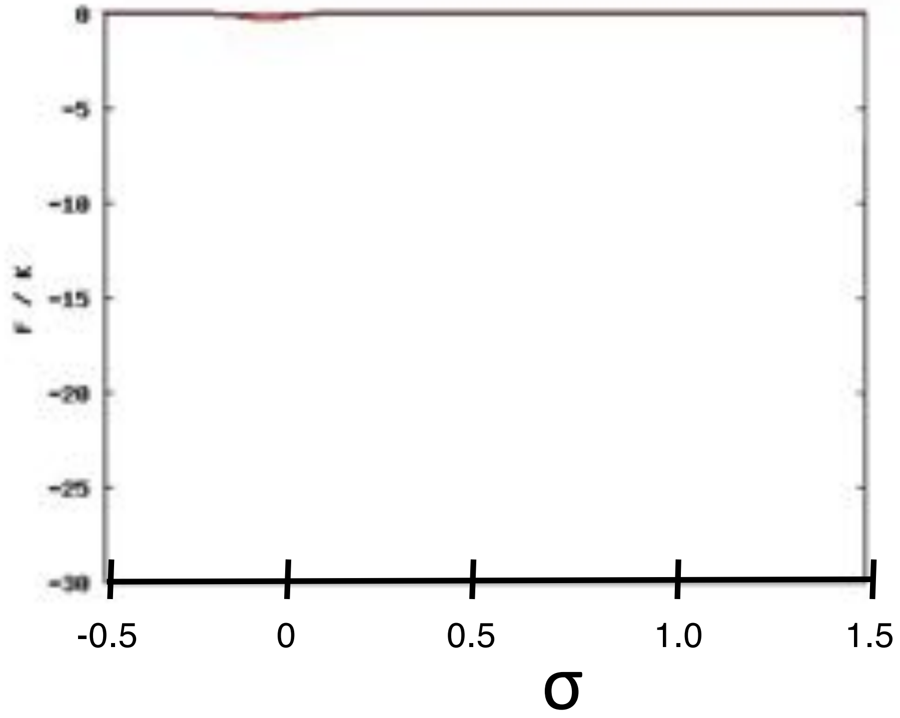
H x 50%

W x 50%

τ x 100







The end

"Mr. Osborne, may I be excused? My brain is full."