COMPUTER SIMULATIONS II

PART I (Kurt Kremer)

Basics of MC and MD

Ensembles

Thermostats

Polymers

COMPUTER SIMULATIONS II

PART II (DF)

Why MD works (?)

MC Cluster moves

Free energy and Phase equilibria

Rare events

AND … Your questions.

But first:

Why do simulations at all?

Why Simulation?

1. Predicting properties of (new) materials
2. Understanding phenomena on a molecular scale
3. Simulating known phenomena **?**

Example: computing the melting point of ice

Why bother?

This works better.

Why Simulation?

1. **Predicting properties of (new) materials**
2. **Understanding phenomena on a molecular scale.**
3. ~~**Reproducing known phenomena.**~~

Testing simulation models by:
Reproducing known phenomena.

Testing approximate theories by:
Computer "experiments"

The limits of Simulation

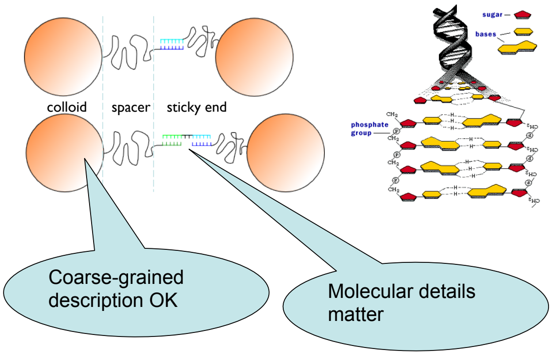
Brute-force simulations can never bridge all the scales between microscopic (nanometers/picoseconds) and macroscopic (cells, humans, planets).

Hence: we need different levels of description ("coarse graining") - and we need input from experiments at many different levels to validate our models.

This is why simulations of Soft Matter are always coarse-grained…

…except when molecular details matter.

Example: DNA-coated colloids



Coarse-grained description OK

Molecular details matter

Why are simulations of Complex Liquids special?

1. Time-scales: long (hydrodynamics, entanglement, …)

2. Interactions: free energies – not just energies

⇒ Algorithms different

Minimal requirements:

a. A `long-enough' simulation should sample the available phase space (MC & MD): `ergodicity'.

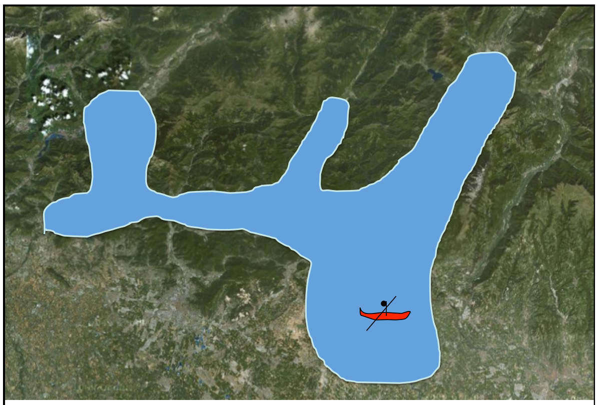
b. The time evolution of the system should be well approximated. (MD)

In practice, ergodicity problems often occur at low temperatures

(glasses, gels, disordered crystals, …)

ANALOGY:

BEFORE GLOBAL WARMING



AFTER (a lot of) GLOBAL WARMING

In Statistical Mechanics language:

“Glassy” energy landscapes.

At low temperatures, breaks up into many “ponds”,

At high temperatures: one “ergodic” lake.

Parallel Tempering

COMBINE “ante-diluvial”and “post-diluvial” in a SINGLE Parallel simulation

In practice:

System 1 at temperature T_1

Boltzmann factor

$$\exp(-\beta_1 U_1(r^N))$$

System 2 at temperature T_2

Boltzmann factor

$$\exp(-\beta_2 U_2(r^N))$$

Total Boltzmann factor

$$\exp(-\beta_1 U_1(r^N))\exp(-\beta_2 U_2(r^N))$$

SWAP move

System 1 at temperature T_2

Boltzmann factor

$$\exp(-\beta_2 U_1(r^N))$$

System 2 at temperature T_1

Boltzmann factor

$$\exp(-\beta_1 U_2(r^N))$$

Total Boltzmann factor

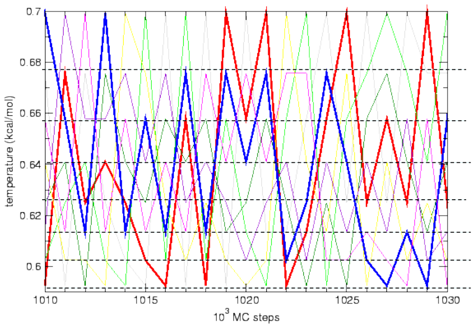
$$\exp(-\beta_2 U_1(r^N))\exp(-\beta_1 U_2(r^N))$$

Ratio

$$\frac{\text{new Boltzmann weight}}{\text{old Boltzmann weight}} = \exp\{-(\beta_1 - \beta_2)[\mathcal{U}_2(\mathbf{r}^N) - \mathcal{U}_1(\mathbf{r}^N)]$$

Systems may swap temperature if their combined Boltzmann factor allows it.

$$\text{acc(swap)} = \min\left(1, \exp\{-(\beta_1 - \beta_2)[\mathcal{U}_2(\mathbf{r}^N) - \mathcal{U}_1(\mathbf{r}^N)]\}\right)$$



The shaky foundations of Molecular Dynamics

…or: why the Verlet algorithm is special.

But first: INTERMEZZO

Stigler’s law of Epynomy states:

No scientific discovery is named after its original discoverer.

and indeed:

Brownian motion was first reported by
Ingen-Housz (or Lucretius?)

"Observe what happens when sunbeams are admitted into a building and shed light on its shadowy places.

You will see a multitude of tiny particles mingling in a multitude of ways... **their dancing is an actual indication of underlying movements of matter that are hidden from our sight... It originates with the atoms that move spontaneously.**

[…] So the movement mounts up from the atoms and gradually emerges to the level of our senses, so that those bodies are in motion that we see in sunbeams, moved by blows that remain invisible.”

Lucretius, 60 BC

The shaky foundations of…

Molecular Dynamics

Molecular Dynamics

The Basis:

$F_i=m_ia_i$

$i=1,2,...,N$

$$X(t+\Delta t) = X(t) + \dot{X}(t)\Delta t + \frac{1}{2!}\ddot{X}(t)\Delta t^2 + \frac{1}{3!}\dddot{X}(t)\Delta t^3 + \frac{1}{4!}\dddot{X}(t)\Delta t^4 + \ldots$$

$$X(t-\Delta t) = X(t) - \dot{X}(t)\Delta t + \frac{1}{2!}\ddot{X}(t)\Delta t^2 - \frac{1}{3!}\dddot{X}(t)\Delta t^3 + \frac{1}{4!}\dddot{X}(t)\Delta t^4 + \ldots$$

+

$$X(t+\Delta t) + X(t-\Delta t) = 2X(t) + \ddot{X}(t)\Delta t^2 + O(\Delta t^4)$$

or

$$\boxed{X(t+\Delta t) \approx 2X(t) - X(t-\Delta t) + \ddot{X}(t)\Delta t^2}$$

Verlet algorithm

Verlet algorithm has attractive features:

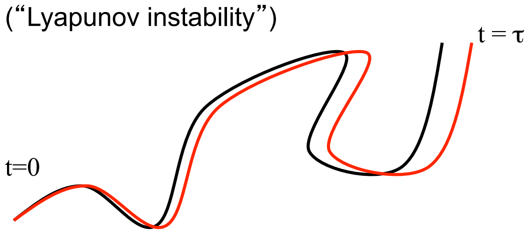
- is time reversible
- does conserve volume in phase space
- does not suffer from energy drift

...but is it a good algorithm?

i.e. does it predict the time evolution of the system correctly???

Dynamics of “well-behaved” classical many-body system is chaotic. Consequence:

Trajectories that differ very slightly in their initial conditions DIVERGE EXPONENTIALLY (“Lyapunov instability”)



If MD cannot predict the time evolution of a system, then

Why should anyone believe Molecular Dynamics simulations ???

We believe that all is well (probably), because of…

The shadow Theorem

For any realistic many-body system, the shadow theorem is merely a hypothesis.

It states that (my words):

Good algorithms generate numerical trajectories that are “close to” a REAL trajectory of the many-body system.

Question:

Does the Verlet algorithm indeed generate “shadow” trajectories?

Take a different look at the problem.

Do not discretize NEWTON’s equation of motion...

...but discretize the ACTION

Intermezzo:

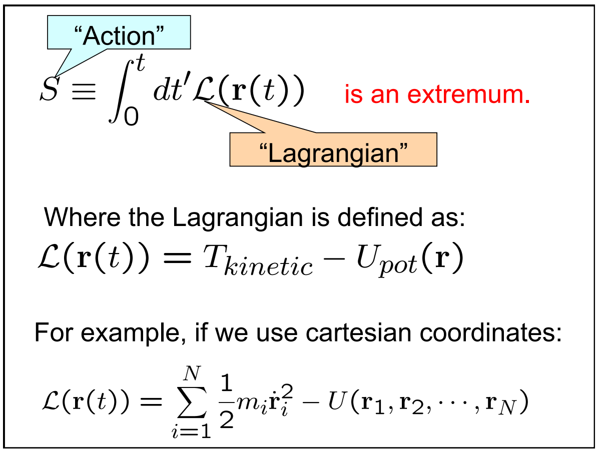
Classical mechanics – the Lagrangian approach.

Newton: $\mathbf{F_i} = m_i \ddot{\mathbf{r}}_i$

Lagrange:

Consider a system that is at a point $\mathbf{r_0}$ at time t=0 and at point $\mathbf{r_t}$ at time t=t, then:

The system follows a trajectory r(t) such that:



"Action"

$$S \equiv \int_0^t dt' \mathcal{L}(\mathbf{r}(t))$$ is an extremum.

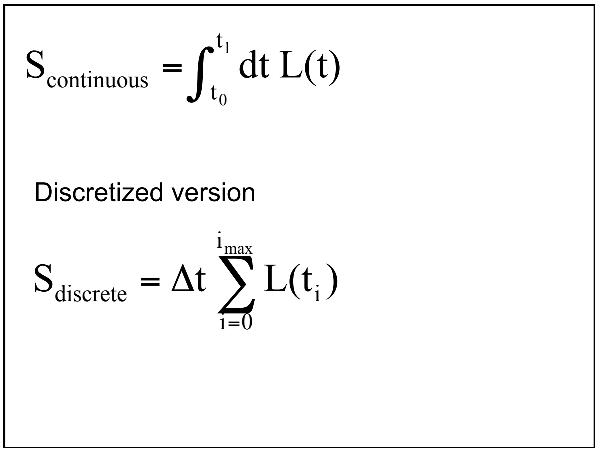
"Lagrangian"

Where the Lagrangian is defined as:

$$\mathcal{L}(\mathbf{r}(t)) = T_{kinetic} - U_{pot}(\mathbf{r})$$

For example, if we use cartesian coordinates:

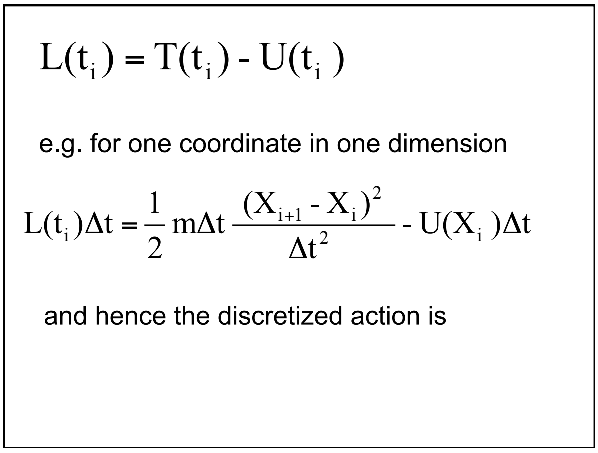
$$\mathcal{L}(\mathbf{r}(t)) = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 - U(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)$$



$$S_{\mathrm{continuous}} = \int_{t_0}^{t_1} dt\, L(t)$$

Discretized version

$$S_{\mathrm{discrete}} = \Delta t \sum_{i=0}^{i_{\mathrm{max}}} L(t_i)$$

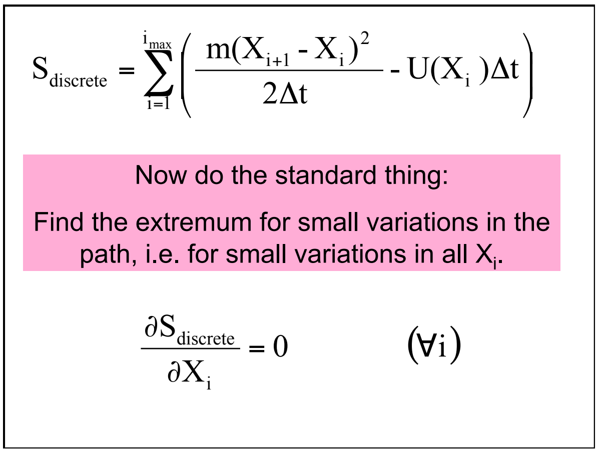


$$L(t_i) = T(t_i) - U(t_i)$$

e.g. for one coordinate in one dimension

$$L(t_i)\Delta t = \frac{1}{2} m \Delta t \frac{(X_{i+1} - X_i)^2}{\Delta t^2} - U(X_i)\Delta t$$

and hence the discretized action is



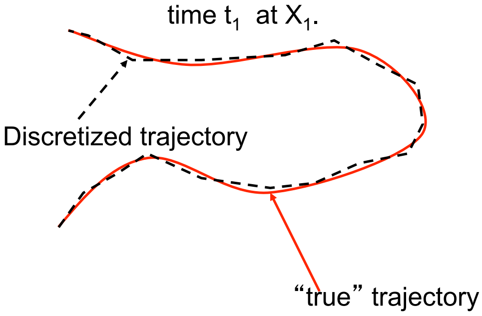
$$S_{\mathrm{discrete}} = \sum_{i=1}^{i_{\mathrm{max}}} \left(\frac{m(X_{i+1} - X_i)^2}{2\Delta t} - U(X_i)\Delta t \right)$$

Now do the standard thing:

Find the extremum for small variations in the path, i.e. for small variations in all X_i.

$$\frac{\partial S_{\mathrm{discrete}}}{\partial X_i} = 0 \qquad (\forall i)$$

This will generate a discretized trajectory that starts at time t_0 at X_0 , and ends at time t_1 at X_1.



$$\frac{\partial S_{discrete}}{\partial X_i} = \frac{\partial}{\partial X_i}\sum_{i=1}^{i_{max}}\left(\frac{m(X_{i+1} - X_i)^2}{2\Delta t} - U(X_i)\Delta t\right)$$

$$\frac{\partial S_{discrete}}{\partial X_i} = \frac{-m(X_{i+1} - X_i) + m(X_i - X_{i-1})}{\Delta t} - \Delta t\frac{\partial U(X_i)}{\partial X_i}$$

And hence:

$$0 = \frac{m}{\Delta t}\left(2X_i - X_{i+1} - X_{i-1} - \frac{\Delta t^2}{m}\frac{\partial U(X_i)}{\partial X_i}\right)$$

$$0 = \left(2X_i - X_{i+1} - X_{i-1} - \frac{\Delta t^2}{m}\frac{\partial U(X_i)}{\partial X_i}\right)$$

REWRITE AS:

$$X_{i+1} = 2X_i - X_{i-1} + \frac{\Delta t^2}{m}F(X_i)$$

VERLET!!!

The Verlet algorithm generates trajectory that satisfies the boundary conditions of a REAL trajectory – both at the beginning and at the endpoint.

Hence, if we are interested in statistical information about the dynamics (e.g. time-correlation functions, transport coefficients, power spectra...)

...then a “good” MD algorithm (e.g. Verlet) is fine.

It is **not** good as a `predictive’ algorithm.

Monte Carlo and cluster moves

$$\mathcal{N}(o)\alpha(o \to n) \times \mathrm{acc}(o \to n) = \mathcal{N}(n)\alpha(n \to o) \times \mathrm{acc}(n \to o)$$

Boltzmann weights

Probability to attempt move

Probability to accept move

Metropolis Monte Carlo:

$\alpha(o \to n) = \alpha(n \to o)$

1. generate trial moves
2. Move if accepted
3. Otherwise, stay where you are

Metropolis, Rosenbluth, Rosenbluth, Teller and Teller:

$$\mathrm{acc}(o \to n) = \min\left(1, \exp\{-\beta[\mathcal{U}(\mathbf{r}'^N) - \mathcal{U}(\mathbf{r}^N)]\}\right)$$

Unsatisfactory?

However: if we do **not** impose

$$\alpha(o \to n) = \alpha(n \to o) \quad \text{then}$$

$$\frac{\mathsf{acc}(o \to n)}{\mathsf{acc}(n \to o)} = \frac{\alpha(n \to o)}{\alpha(o \to n)} \mathsf{exp}\{-\beta[\mathcal{U}(n) - \mathcal{U}(o)]\}.$$

In particular, if:

$$\frac{\alpha(n \to o)}{\alpha(o \to n)} = \mathsf{exp}\{-\beta[\mathcal{U}(o) - \mathcal{U}(n)]\}.$$

Then

$$\frac{\mathsf{acc}(o \to n)}{\mathsf{acc}(n \to o)} = 1 \qquad \text{(100\% acceptance)}$$

100% acceptance can be achieved in special cases:
e.g. Swendsen-Wang algorithm

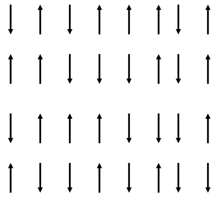
Discrete spin models (Potts, Ising).

Illustration: 2D Ising model:

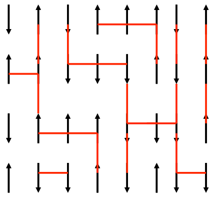
Parallel nearest neighbor spins: energy **–J**

Anti-parallel nearest neighbor spins: energy **+J**

$$U = -J \sum_{i,j} s_i s_j$$



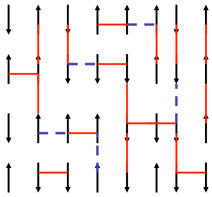
Snapshot: some neighbors are parallel,
others anti-parallel



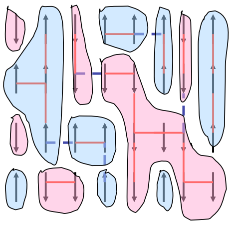
Count number of bonds between parallel neighbors: $\mathbf{N_p}$

Number of bonds between anti-parallel neigbors is: $\mathbf{N_a}$

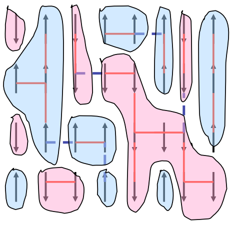
Total energy: $\mathbf{U = (N_a - N_p)\, J}$



Now, make “bonds”. Bonds only form between parallel neighbors. The probability to have a bond (red line) between parallel neighbors is **p** (as yet undetermined). With a probability **1-p**, parallel neighbors are not connected (blue dashed line).

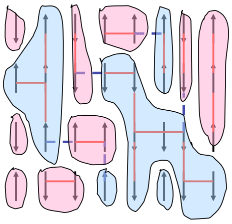


Form clusters of all spins that are connected by bonds. Some clusters are all “spin up” others are all “spin down”. Let us denote the number of clusters by **M**.



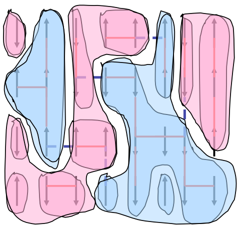
The probability to generate a particular cluster structure where there are $\mathbf{n_c}$ bonds between $\mathbf{N_p}$ pairs of parallel neighbors is:

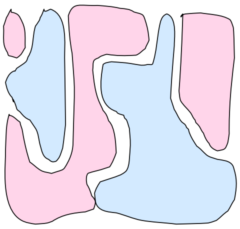
$$P_{gen} = p^{n_c}(1-p)^{N_p-n_c}$$



Now randomly flip clusters. This yields a new cluster configuration with probability $\mathbf{P_{(flip)}=(1/2)^M}$.

Then reconnect parallel spins





New cluster structure!

Now make it into a Monte Carlo algorithm:

$$P_o P_{clus}(o) P_{flip}(M) P_{acc}(o \to n)$$

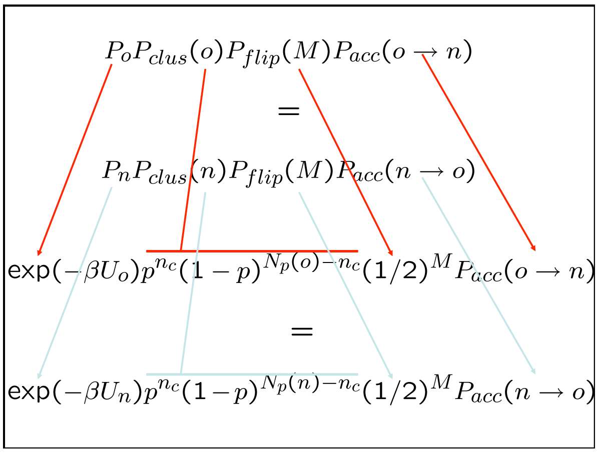
$$=$$

$$P_n P_{clus}(n) P_{flip}(M) P_{acc}(n \to o)$$

$$\exp(-\beta U_o) p^{n_c} (1-p)^{N_p(o)-n_c} (1/2)^M P_{acc}(o \to n)$$

$$=$$

$$\exp(-\beta U_n) p^{n_c} (1-p)^{N_p(n)-n_c} (1/2)^M P_{acc}(n \to o)$$



$$P_o P_{clus}(o) P_{flip}(M) P_{acc}(o \to n)$$

$$=$$

$$P_n P_{clus}(n) P_{flip}(M) P_{acc}(n \to o)$$

$$\exp(-\beta U_o) p^{n_c} (1-p)^{N_p(o)-n_c} (1/2)^M P_{acc}(o \to n)$$

$$=$$

$$\exp(-\beta U_n) p^{n_c} (1-p)^{N_p(n)-n_c} (1/2)^M P_{acc}(n \to o)$$

Moreover, we want 100% acceptance, i.e.:

$P_{acc}(o \rightarrow n) = P_{acc}(n \rightarrow o) = 1$

$$\exp(-\beta U_o)\cancel{p^{n_c}}(1-p)^{N_p(o)-\cancel{n_c}}(\cancel{1/2})^M P_{acc}(o \to n)$$

$$=$$

$$\exp(-\beta U_n)\cancel{p^{n_c}}(1-p)^{N_p(n)-\cancel{n_c}}(\cancel{1/2})^M P_{acc}(n \to o)$$

Hence:

$$\exp(-\beta U_o)(1-p)^{N_p(o)} = \exp(-\beta U_n)(1-p)^{N_p(n)}$$

$$\exp(\beta(U_n - U_o)) = (1-p)^{N_p(n)-N_p(o)}$$

But remember:

$$U_n - U_o = J(N_a(n) - N_p(n)) - J(N_a(o) - N_p(o))$$

or

$$\Delta U = J(\Delta N_a - \Delta N_p)$$

But: $\Delta N_a = -\Delta N_p$

and therefore

$$\Delta U = -2J\Delta N_p$$

$$\exp(\beta(U_n - U_o)) = \exp(-2\beta J(N_p(n) - N_p(o)))$$

Combining this with:

$$\exp(\beta(U_n - U_o)) = (1-p)^{N_p(n)-N_p(o)}$$

we obtain:

$$p = 1 - \exp(-2\beta J)$$

100% acceptance!!!

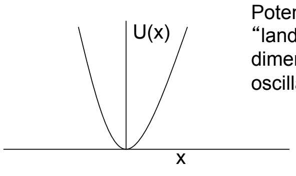
Free energy

---

1. Free-energy landscapes?
2. Phase coexistence

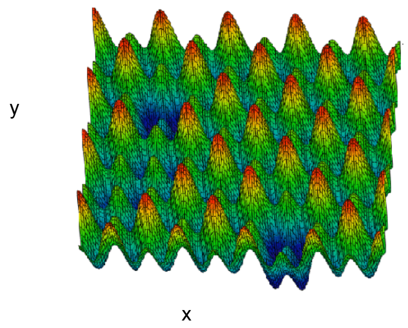
First question:
What **IS** a free-energy landscape?

Simpler question:
What is an energy landscape?



Potential-energy
“landscape” of a 1-dimensional harmonic
oscillator

2-dimensional potential energy landscape (e.g. surface potential experienced by an adsorbate atom)

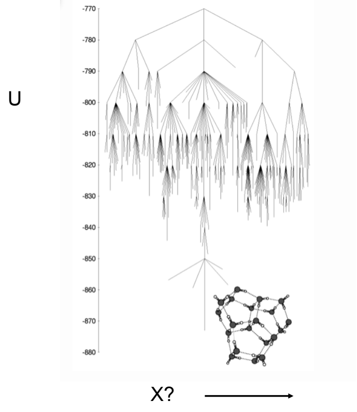


General potential energy landscape:

$$U(\mathbf{r}^N) \text{ with } \mathbf{r}^N \equiv (x_1, y_1, z_1, \cdots, x_i, y_i, z_i \cdots x_N, y_N, z_N)$$

High dimensional - not easy to visualise.

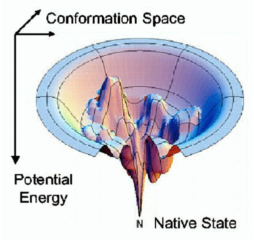
Visual aid: “disconnectivity graphs”



“tree” of energy minima.

Topology - not distance

Free energy landscape?



But what are the landscape coordinates?

In order to define a free energy it is necessary to specify the coordinates of the landscape.

Other coordinates => other free-energy landscape"

Statistical mechanics: Boltzmann weight.

$$P(r^N) = \frac{e^{-U(\mathbf{r}^N)/kT}}{Z}$$

What is the probability that the center of mass of the system is at a coordinate X?

$$P(X) = \frac{\int dx^N e^{-U(x^N)/kT}\delta(X - N^{-1}\sum_i x_i)}{Z}$$

We now *define* the free energy associated with center-of-mass coordinate X as:

$$e^{-F(X)/kT} \equiv \int dx^N e^{-U(x^N)/kT}\delta(X - N^{-1}\sum_i x_i)$$

The free energy is to the "collective" coordinate X, what the potential energy is to the individual coordinates.

In general, there may be several coordinates, X, Y, Z etc.

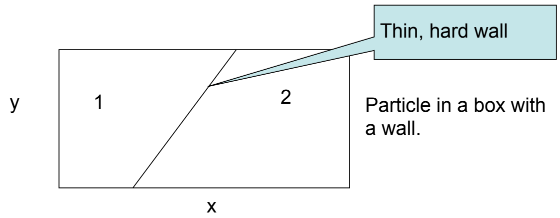
They may be complicated functions of $\mathbf{r}^N$, and they may be discrete.
e.g.
X = radius of gyration of a protein
Y = number of native contacts

One message:

There is no such thing as **the** free energy landscape of a system.

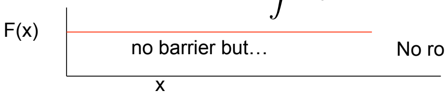
We can only define F(X,Y,…) *after* choosing the relevant coordinates X,Y,…

Note: there may be a landscape but not a road…



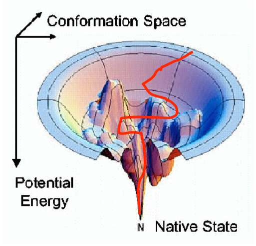
Particle in a box with a wall.

$$e^{-F(x)/k_BT} = \int dy\; e^{-U(x,y)/k_BT}$$



No road!

A "funnel" may, or may not contain a "road".



Usually, it does.

Disconnectivity graphs **always** contain a road (but no distance).

When does free energy matter?

a. Phase transitions
b. Solubility/aggregation/reactions
c. Substrate binding
d. Conformational changes/folding
e. … and much more

Why are free-energy calculations a separate topic?

Pressure
Temperature
Heat capacity
…
Observables: can be "sampled"

Free energy: **not**

Example:

predicting phase coexistence

Free-energy “coordinate” is discrete.

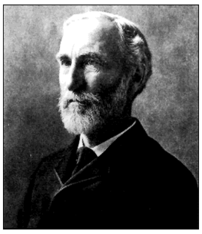
Two values:
1 if system is in phase 1
2 if system is in phase 2

Condition for phase coexistence in a one-component system:

$$T_1 = T_2$$

$$P_1 = P_2$$

$$\mu_1 = \mu_2$$



J.Willard Gibbs

But if we know F we can compute μ and vice versa:

$$\mu = \frac{G}{N} = \frac{F + PV}{N}$$

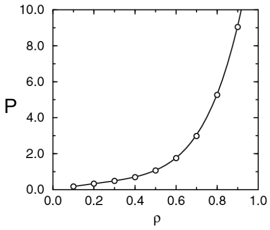
How to obtain F or μ from simulation?

1. “normal” thermodynamic integration
2. “artificial” thermodynamic integration
3. “particle-insertion” method
4. …

How are free energies measured experimentally?

$$\frac{\partial F}{\partial V} = -P$$

$$\frac{\partial F/T}{\partial 1/T} = E$$



$$F(V) = F(V_0) + \int_{V_0}^{V} (-P)dV$$

Then take the limit $V \to \infty$

Not so convenient because of divergences.

Better:

$$F_{excess}(V) = F(V) - F_{id.gas}(V)$$

$$= F_{excess}(V_0) + \int_{V_0}^{V} (-P_{excess})dV$$

⇒ 0, as $V_0 \Rightarrow \infty$

This approach works if we can integrate from a known reference state, e.g.

Ideal gas (“T=∞”),

Harmonic crystal (“T=0”)

Otherwise: use “artificial” thermodynamic integration (Kirkwood)

Suppose we know F(N,V,T) for a system with a simple potential energy function U_0: $F_0(N,V,T)$.

We wish to know $F_1(N,V,T)$ for a system with a “complex” potential energy function U_1.

Consider a system with a mixed potential energy function $(1-\lambda)U_0 + \lambda U_1$:

Then (2-line derivation skipped):

$$\frac{\partial F_\lambda(N,V,T)}{\partial \lambda} = \langle U_1 - U_0 \rangle_\lambda$$

And therefore

$$F_1(N,V,T) = F_0(N,V,T) + \int_0^1 \langle U_1 - U_0 \rangle_\lambda d\lambda$$

Example:

Phase diagram of hard rod-like colloids.

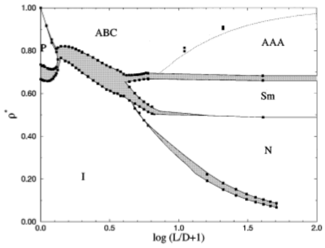


FIG. 2. Summary of the phase diagram of hard spherocylinders with L/D between 0 and 100. In order to give equal emphasis to all parts of the phase diagram, we have plotted ρ^* as a function of $\log(L/D+1)$. The dashed line is a crude estimate for the first order AAA-ABC transition as given in Eq. (52).

Particle insertion method to compute chemical potential

Thermo:

dF = -PdV -SdT +μdN

Using F=- k_BT ln Q(N,V,T) we can write:

$$\begin{aligned}\mu &= \left(\frac{\partial F}{\partial N}\right)_{V,T}\\ &= -k_B T \lim_{\Delta N\to 0}\frac{\ln Q(N+\Delta N,V,T)-\ln Q(N,V,T)}{\Delta N}\\ &= -k_B T \ln\left(\frac{Q(N+1,V,T)}{Q(N,V,T)}\right)_{V,T}\end{aligned}$$

For N >> 1

Example: ideal gas

$$Q(N,V,T)=\frac{V^N}{\Lambda^{3N}N!}$$

$$\begin{aligned}\mu &= \lim_{N\to\infty} -k_B T \ln\left[\frac{\frac{V^{N+1}}{\Lambda^{3(N+1)}(N+1)!}}{\frac{V^N}{\Lambda^{3N}N!}}\right]\\ &= -k_B T \ln\frac{V}{\Lambda^3(N+1)}\\ &\approx k_B T \ln\left(\Lambda^3\rho\right)\end{aligned}$$

with ρ ≡ (N/V)

The ideal-gas result, we had obtained before.
Now, with interacting particles:

$$\begin{aligned}\mu &= \lim_{N\to\infty} -k_BT\ln\left[\frac{Q(N+1,V,T)}{Q(N,V,T)}\right]\\ &\equiv \mu^{\rm id}+\mu^{\rm excess}\end{aligned}$$

We can there write the "excess" part of the chemical potential as:

$$\begin{aligned}\mu^{\rm ex} &= -k_BT\ln\left[\frac{Q(N+1,V,T)}{Q(N,V,T)} - \ln\left(\frac{V}{\Lambda^3(N+1)}\right)\right]\\ &= -k_BT\ln\left[\frac{Q(N+1,V,T)\Lambda^3(N+1)}{VQ(N,V,T)}\right]\end{aligned}$$

Now we use the explicit expression for Q(N,V,T):

$$Q(N,V,T) = \frac{1}{\Lambda^{3N}N!}\int d\mathbf{r}^N e^{-\beta U(\mathbf{r}^N)}$$

$$\mu^{\rm ex} = -k_BT\ln\left(\frac{\int d\mathbf{r}^{N+1}e^{-\beta U(\mathbf{r}^{N+1})}}{V\int d\mathbf{r}^N e^{-\beta U(\mathbf{r}^N)}}\right)$$

We define:

$$\Delta U(\mathbf{r}^N;\mathbf{r}_{N+1}) \equiv U(\mathbf{r}^{N+1}) - U(\mathbf{r}^N)$$

Then

$$\begin{aligned}\mu^{\rm ex} &= -k_BT\ln\left(\frac{\int d\mathbf{r}^{N+1}e^{-\beta[U(\mathbf{r}^N)+\Delta U(\mathbf{r}^N;\mathbf{r}_{N+1})]}}{\int d\mathbf{r}^N e^{-\beta U(\mathbf{r}^N)}}\right)\\ &= -k_BT\ln\left(\frac{1}{V}\int d\mathbf{r}_{N+1}\frac{\int d\mathbf{r}^N e^{-\beta[U(\mathbf{r}^N)+\Delta U(\mathbf{r}^N;\mathbf{r}_{N+1})]}}{\int d\mathbf{r}^N e^{-\beta U(\mathbf{r}^N)}}\right)\end{aligned}$$

$$\begin{aligned}\mu^{\rm ex} &= -k_BT\ln\left(\frac{1}{V}\int d\mathbf{r}_{N+1}\frac{\int d\mathbf{r}^N e^{-\beta U(\mathbf{r}^N)}e^{-\beta\Delta U(\mathbf{r}^N;\mathbf{r}_{N+1})}}{\int d\mathbf{r}^N e^{-\beta U(\mathbf{r}^N)}}\right)\\ &= -k_BT\ln\frac{1}{V}\int d\mathbf{r}_{N+1}\left\langle e^{-\beta\Delta U(\mathbf{r}^N;\mathbf{r}_{N+1})}\right\rangle_N\end{aligned}$$

For a homogeneous system, the average does not depend on the position $\mathbf{r}_{N+1}$. Hence

$$\beta\mu^{\rm ex} = -\ln\left\langle e^{-\beta\Delta U}\right\rangle_N$$

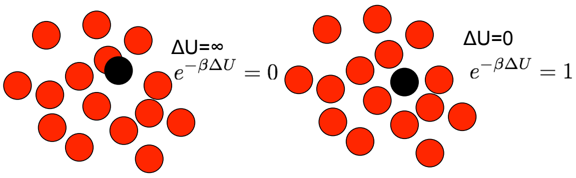
In words: the excess chemical potential is equal to minus the logarithm of the average Boltzmann factor associated with the **random insertion** of an additional particle N+1 in an N-particle system.

This method to evaluate μ is commonly known as **Widom**'s "particle-insertion method".

EASY TO ADD ON TO ANY MD/MC PROGRAMME.

Fails at high densities.

Example: excess chemical potential of "hard" particles



Particle insertion continued….

$$\mu \approx \left(\frac{F(N+1,V,T)-F(N,V,T)}{(N+1)-N}\right)$$

therefore

$$\mu \approx -kT \ln \frac{Q(N+1,V,T)}{Q(N,V,T)}$$

But also

$$\mu \approx +kT \ln \frac{Q(N,V,T)}{Q(N+1,V,T)}$$

$$\frac{Q(N,V,T)}{Q(N+1,V,T)} = \frac{(N+1)\Lambda^3}{V}\langle \exp(+\beta\Delta U(\mathbf{r}_{N+1},\mathbf{r}^N))\rangle_{N+1}$$

And therefore

$$\mu = \mu_{id.gas} + kT\ln\langle\exp(+\beta\Delta U(\mathbf{r}_{N+1},\mathbf{r}^N))\rangle_{N+1}$$

~~Interpretation:~~

1. ~~Evaluate ΔU for a random **REMOVAL** of a molecule in a system containing N+1 molecule.~~
2. ~~Compute~~ $\cancel{\exp(+\beta\Delta U)}$
3. ~~Repeat M times and compute the average "Boltzmann factor"~~
4. ~~Then~~

$$\cancel{\mu_{excess} = +kT \ln < \exp(+\beta\Delta U) >}$$

To some, this is very disappointing…

What is wrong?

$\exp(+\beta\Delta U)$ is not bounded. The average that we compute can be dominated by INFINITE contributions from points that are NEVER sampled.

What to do?

Consider:

$$p_0(\Delta U) \equiv \frac{\int \exp(-\beta U_N)\delta(\Delta U - U_{N+1} + U_N)}{\int \exp(-\beta U_N)}$$

And also consider the distribution

$$p_1(\Delta U) \equiv \frac{\int \exp(-\beta U_{N+1})\delta(\Delta U - U_{N+1} + U_N)}{\int \exp(-\beta U_{N+1})}$$

p_0 and p_1 are related:

$$p_1(\Delta U) = \frac{\int \exp(-\beta(U_N + \Delta U))\delta(\Delta U - U_{N+1} + U_N)}{Q_{N+1}}$$

$$= \exp(-\beta\Delta U)\frac{\int \exp(-\beta U_N)\delta(\Delta U - U_{N+1} + U_N)}{Q_{N+1}}$$

$$= \exp(-\beta\Delta U)\frac{Q_N}{Q_{N+1}}\frac{\int \exp(-\beta U_N)\delta(\Delta U - U_{N+1} + U_N)}{Q_N}$$

so

$$p_1(\Delta U) = \exp(-\beta\Delta U)\frac{Q_N}{Q_{N+1}}p_0(\Delta U)$$

but $\dfrac{Q_N}{Q_{N+1}} = \exp(+\beta\mu)$ hence

$$p_1(\Delta U) = \exp(-\beta\Delta U)\exp(\beta\mu)p_0(\Delta U)$$

Rewrite:

$$\ln(p_1(\Delta U)) = -\beta\Delta U + \beta\mu + \ln(p_0(\Delta U))$$

define

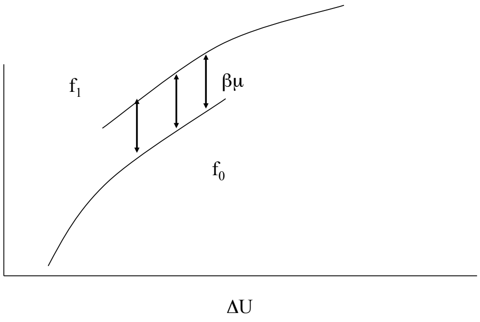
$$f_1(\Delta U) \equiv \ln(p_1(\Delta U)) + \beta\Delta U/2$$

$$f_0(\Delta U) \equiv \ln(p_0(\Delta U)) - \beta\Delta U/2$$

Then:

$$f_1(\Delta U) = f_0(\Delta U) + \beta\mu$$

For **ALL** ΔU



Diagnostic test:

If the distributions do not overlap, **don't trust your results for μ**

(similar expression for path sampling ⇔ van den Broeck)

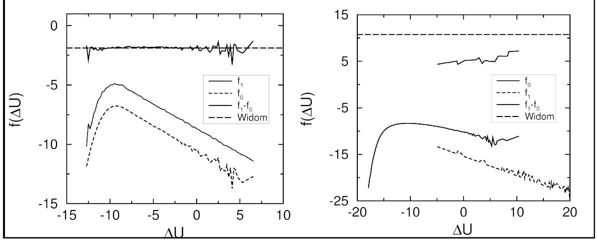
Chemical potential

System 0: *N-1, V,T, U* + 1 ideal gas System 1: *N, V,T, U*

$$\Delta\beta F = \beta F_1 - \beta F_0 \equiv \beta\mu^{ex} \qquad \Delta U = U_1 - U_0$$

System 0: test particle energy System 1: real particle energy

$$\beta\mu^{ex} = f_1(\Delta U) - f_0(\Delta U)$$



Does it work for hard spheres?

consider ΔU=0

$$f_1(0) = f_0(0) + \beta\mu$$

$$f_1(0) = \ln(1) + (constant)$$

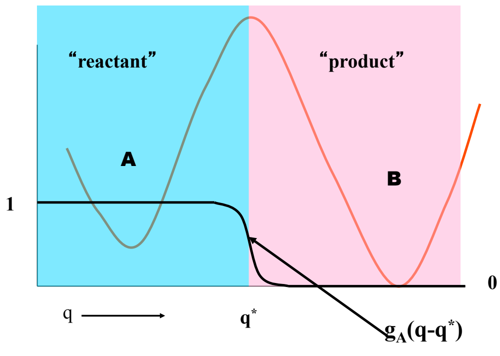
$$f_0(0) = \ln(P_{acc}) + (constant)$$

$$\beta\mu = -\ln(P_{acc})$$

Rare events: crossing free energy barriers.

There exist many techniques to study barrier crossings (Bennett-Chandler, Transition-Path Sampling, String Method, Forward Flux sampling, Milestoning etc etc)

I will discuss two:
1. Bennett-Chandler
2. Forward-flux sampling



$$\mathcal{H} = \mathcal{H}_0 - \epsilon g_A(q - q^*)$$

Common (not best) choice for $g_A(q\text{-}q^*)$:

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

If we normalize the total concentration to 1, then $\langle c_A\rangle = \langle g_A\rangle$, and hence

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

Linear response theory (next slide)

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

$$\frac{\partial \Delta c_A}{\partial \epsilon} = \beta\left[\langle g_A\rangle_0 \left(1 - \langle g_A\rangle_0\right)\right]\} = \beta \langle c_A\rangle \langle c_B\rangle$$

Because:

$$\langle g_B\rangle_0 = \langle (1 - g_A)\rangle_0 = \langle c_B\rangle_0$$

Intermezzo: linear response theory (part).

Consider the response of an observable A due to an external field f_B that couples to an observable B:

$$H = H_0 - f_B B$$

For simplicity, assume that $\langle A\rangle_0 = \langle B\rangle_0 = 0$

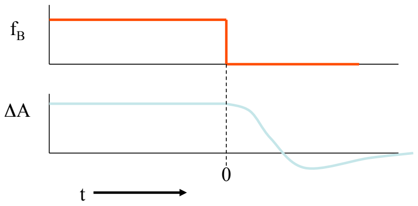
$$\langle \Delta A\rangle_{f_B} = \frac{\int \exp[-\beta(H_0 - f_B B)]A}{\int \exp[-\beta(H_0 - f_B B)]}$$

For small f_B we can linearize:

$$\langle \Delta A\rangle \approx \beta f_B \frac{\int \exp[-\beta H_0]BA}{\int \exp[-\beta H_0]}$$

Hence $\langle \Delta A\rangle \approx \beta f_B \langle BA\rangle_0$

Now consider a weak field that is switched off at t=0.



Using exactly the same reasoning as in the static case, we find:

$$\langle \Delta A\rangle(t) = \beta f_B \langle BA(t)\rangle_0$$

In the present case,

$$A = B = g_A(q - q^*)$$

Time dependent linear response theory:

$$\Delta c_A(t) = \beta\epsilon \langle \Delta g_A(0)\Delta g_A(t)\rangle$$

We eliminate the external field, using

$$\Delta c_A(0) \approx \epsilon\beta \langle c_A\rangle \langle c_B\rangle$$

then

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

The macroscopic equation was:

$$\Delta c_A(t) = \Delta c_A(0)\,\mathrm{exp}(-t/\tau_R)$$

And therefore

$$\mathrm{exp}(-t/\tau_R) = \frac{\langle \Delta g_A(0)\Delta g_A(t)\rangle}{\langle c_A\rangle \langle c_B\rangle}$$

More convenient: consider at time derivative

$$-\tau_R^{-1}\,\mathrm{exp}(-t/\tau_R) = \frac{\langle g_A(0)\dot{g}_A(t)\rangle}{\langle c_A\rangle \langle c_B\rangle}$$

$$= -\frac{\langle \dot{g}_A(0) g_A(t)\rangle}{\langle c_A\rangle \langle c_B\rangle}$$

$$\tau_R^{-1} = \frac{\langle \dot{g}_A(0) g_A(t)\rangle}{\langle c_A\rangle \langle c_B\rangle}$$

$$k_{A\to B}(t) = \frac{\langle \dot{g}_A(0) g_A(t)\rangle}{\langle c_A\rangle}$$

$$\dot{g}_A(q-q^*) = \dot{q}\frac{\partial g_A(q-q^*)}{\partial q}$$

using

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

We get

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

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

This can be rewritten as:

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

$$\times \frac{\langle \delta(q^* - q)\rangle}{\langle \theta(q^* - q)\rangle}$$

Now the second term on the right is simply equal to

$$P(q^*) \equiv \frac{\int \mathrm{dr}\ \exp(-\beta \mathcal{U})\delta(q^* - q)}{\int \mathrm{dr}\ \exp(-\beta \mathcal{U})\theta(q^* - q)}.$$

For the first term on the right we have:

Using statistical mechanics *(0.5 hour)*, we can derive an expression for the barrier crossing rate:

$$k_{A\to B} = \langle \dot{q}(0)\theta(q(t) - q^*)\rangle_{q(0)=q^*}\, P(q^*)$$

Kinetic prefactor

$$\boxed{\exp(-\beta G(q^*))}$$

$$\boxed{k_{A\to B} = \langle \dot{q}(0)\theta(q(t) - q^*)\rangle_{q(0)=q^*}\, P(q^*)}$$

In words:

Rate= equilibrium probability to find system at the top of the barrier, multiplied by

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

The "cross-correlation" of the velocity in the direction of the reaction coordinate at t=0 and the probability that the system has crossed the barrier at time t.

Computing free-energy landscapes

Suppose we have an "reaction coordinate" $Q(r^N)$.

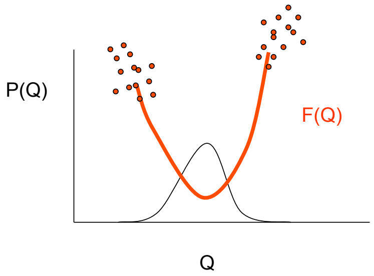
The probability that $Q(r^N)$ has a value Q is given by

$$P(Q_0) = \frac{\int \exp(-\beta U(r^N))\delta(Q(r^N) - Q_0)}{\int \exp(-\beta U(r^N))}$$

From this probability we can derive the variation of the free energy with Q:

$$F(Q) = -kT \ln(P(Q)) + \text{constant}$$

This indicates a route to determine F(Q): just make a histogram of the spontaneous fluctuations of Q.



Problem: F(Q) is very noisy, except near its minimum.

Application of biased MC simulation to determine $P_w(Q)$

$$P_w(Q_0) = \frac{\int \exp[-\beta(U(r^N) + w(Q(r^N))]\delta(Q(r^N) - Q_0)}{\int \exp[-\beta(U(r^N) + w(Q(r^N))]}$$

But this we can rewrite as

$$P_w(Q_0) = \exp(-\beta w(Q_0)) \frac{\int \exp[-\beta U(r^N)]\delta(Q(r^N) - Q_0)}{\int \exp[-\beta(U(r^N) + w(Q(r^N))]}$$

Clearly,

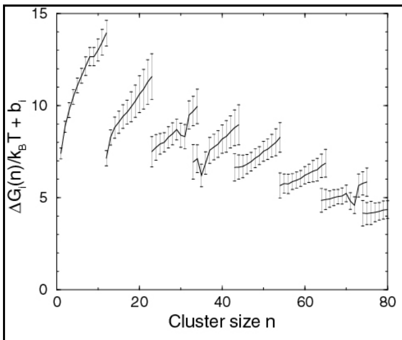
$$P_w(Q_0) = c \exp(-\beta w(Q_0)) P(Q_0)$$

or

$$-kT \ln(P(Q_0)) = -kT \ln(P_w(Q_0)) + w(Q_0) - kT \ln c$$

$$F(Q_0) = -kT \ln(P_w(Q_0)) + w(Q_0) - kT \ln c$$

We can choose the bias w(Q) such that any desired range of Q-values is sampled. And we can correct for the bias.

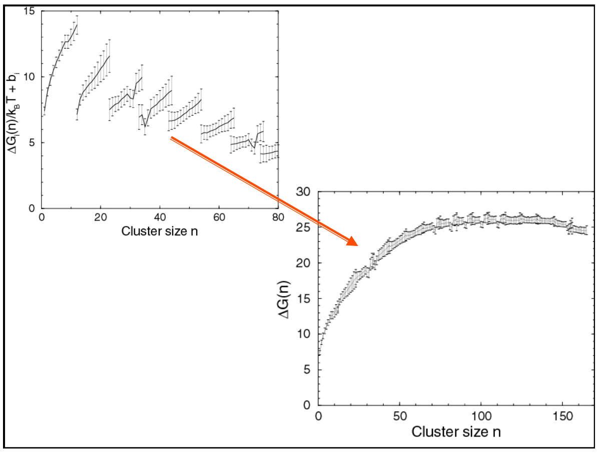


The only remaining problem is that the different parts of F(Q) are shifted with respect to each other.

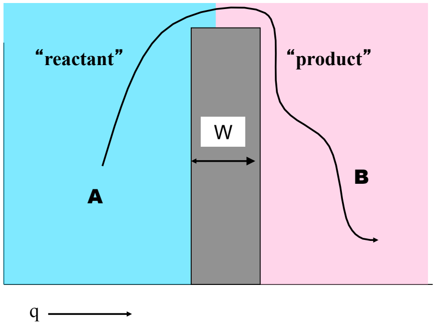
Solutions:

1. Fit to a single curve (not very elegant, but effective)
2. Use Ferrenberg-Swendsen scheme to combine different parts of the histogram (more elegant, but more sensitive to noise)

Result:



Simulating rare events: the diffusive limit…



How do concentration fluctuations decay?

$$\Delta c_A(t) = \Delta c_A(0)\,\exp(-t/\tau_R)$$

If the concentration increase on the left side of the barrier is $\Delta C_A(0)$, then the concentration increase on the top of the barrier is, on the left side:

$$\Delta c_A^* = \Delta c_A \exp(-\beta \Delta G)$$

And, on the right side: 0 ("absorbing boundary")

The diffusive flux across the barrier is:

$$j_A = D(\Delta c_A^*/W)$$

Therefore

$$\Delta \dot{c}_A(t) = -(D/W)\Delta c_A(t)\exp(-\beta\Delta G)$$

The solution is

$$\Delta c_A(t) = \Delta c_A(0)\exp(-t/\tau_R)$$

with

$$\frac{1}{\tau_R} = \frac{D}{W}\exp(-\beta\Delta G)$$

In practice, the expression is only slightly more complex.

To compute the rate of an activated process, we must compute:

1. The free energy barrier (ΔG)
2. The diffusion constant at the top of the barrier (D)
3. The “effective” width of the barrier (for a square barrier, this is W – in other cases it is a bit different)

SOLUTION 2:

Use “path sampling” techniques

(transition path sampling, string method etc etc etc.)

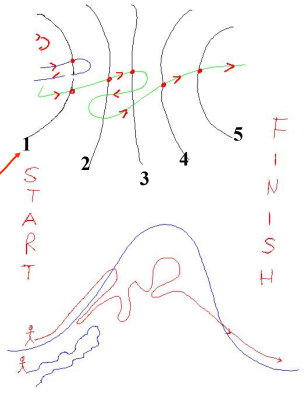
One example: Forward-flux sampling,

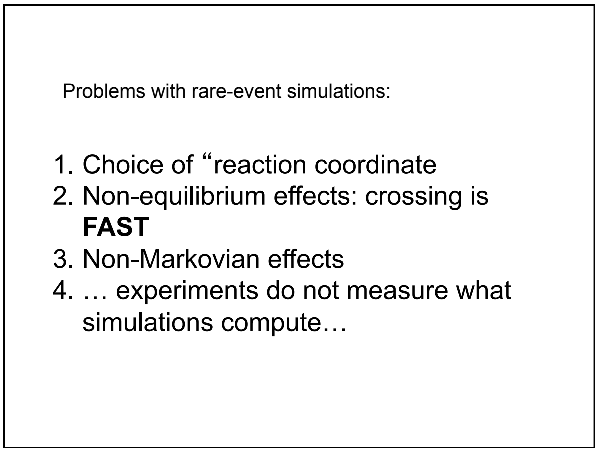
(R.J. Allen et al. Phys. Rev. Lett. 94, 18104(2005))

is that they can be used for NON-EQUILIBRIUM systems.

Basic idea:

1. Define a “reaction” coordinate that distinguishes START from FINISH.
2. Compute the rate Γ_0 (low, but not very low) to reach surface 1.
3. Generate many trajectories from first crossing points and compute the fraction P(12) that reach **2** (without first returning to **START**).
4. **Repeat for all subsequent planes - and for many starting points.**
5. **Rate=** Γ_0 P(12)P(23)..P(5F)





Problems with rare-event simulations:

1. Choice of “reaction coordinate
2. Non-equilibrium effects: crossing is **FAST**
3. Non-Markovian effects
4. … experiments do not measure what simulations compute…