Other ensembles?

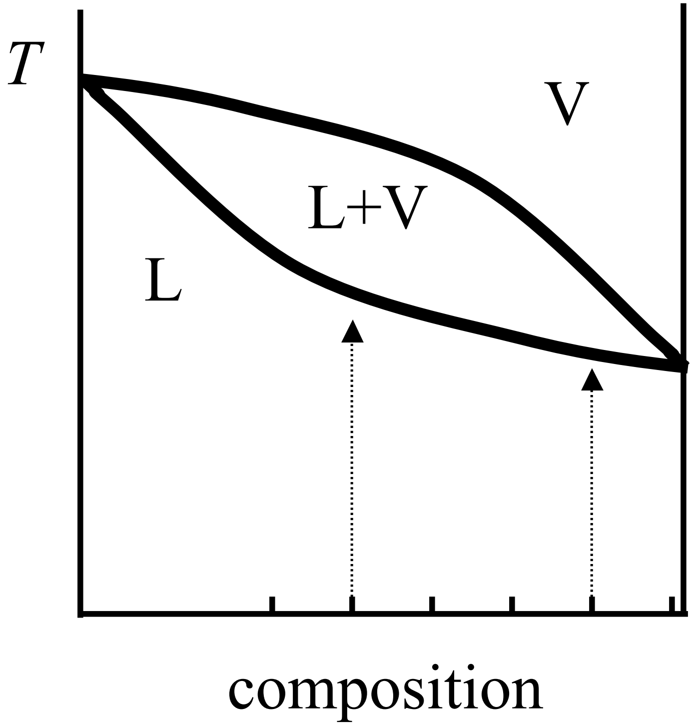
In the thermodynamic limit the thermodynamic properties are independent of the ensemble: so buy a bigger computer …

However, it is most of the times much better to think and to carefully select an appropriate ensemble.

For this it is important to know how to simulate in the various ensembles.

But for doing this wee need to know the Statistical Thermodynamics of the various ensembles.

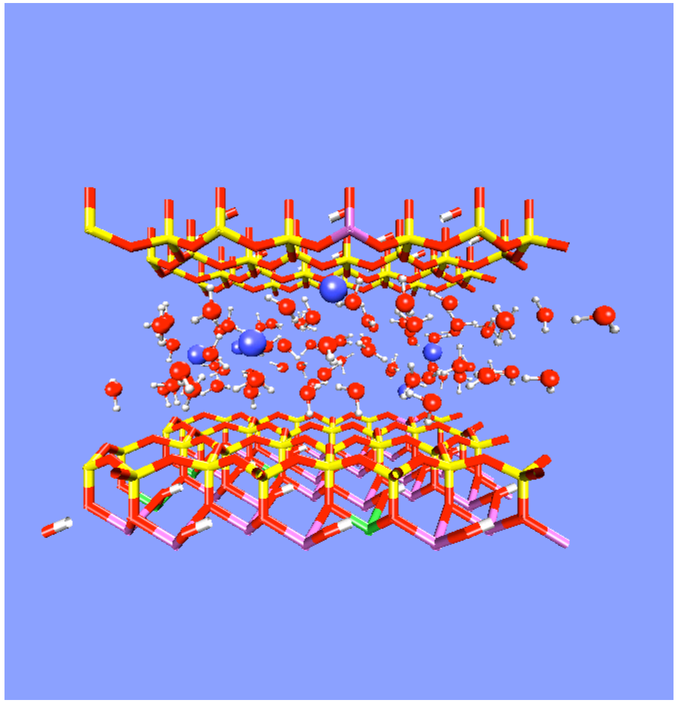
Example (1): vapour-liquid equilibrium mixture



Measure the composition of the coexisting vapour and liquid phases if we start with a homogeneous liquid of two different compositions:

- How to mimic this with the *N,V,T* ensemble?
- What is a better ensemble?

Example (2): adsorption of water in porous medium



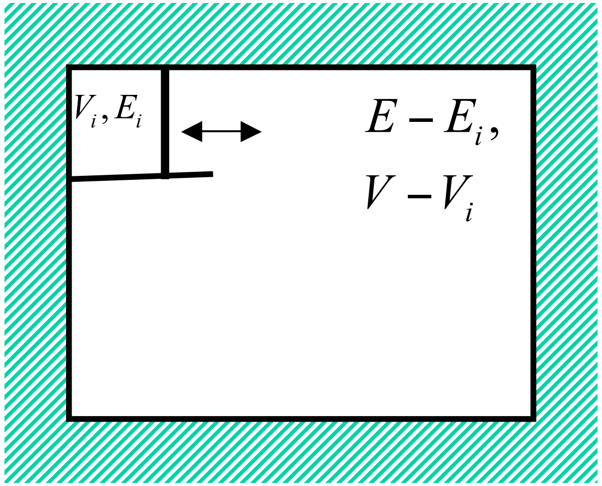
Clay layers can swell upon adsorption of water:

- How to mimic this in the N, V, T ensemble?
- What is a better ensemble to use?

Ensembles

- Micro-canonical ensemble: E,V,N
- Canonical ensemble: T,V,N
- Constant pressure ensemble: T,P,N
- Grand-canonical ensemble: T,V,μ

Constant pressure ensemble: constant N,P,T



Consider a small system that can exchange volume and energy with a large reservoir

$$\ln \Omega\left(V - V_{i,} E - E_i\right) = \ln \Omega\left(V, E\right) - \left(\frac{\partial \ln \Omega}{\partial E}\right)_V E_i - \left(\frac{\partial \ln \Omega}{\partial V}\right)_E V_i + \cdots$$

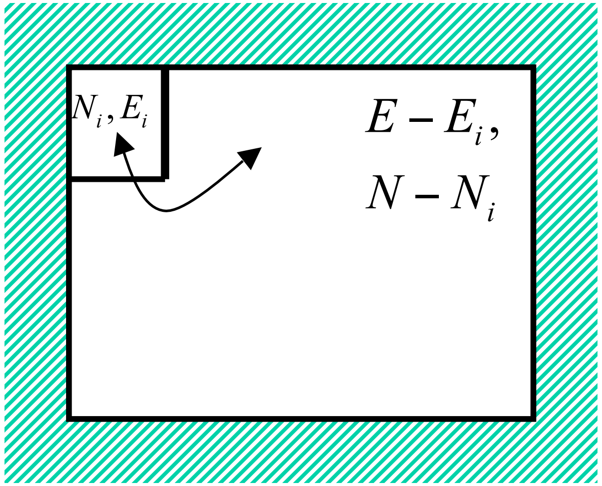
$$\ln \frac{\Omega\left(E - E_i, V - V_i\right)}{\Omega\left(E, V\right)} = -\frac{E_i}{k_B T} - \frac{pV_i}{k_B T}$$

Hence, the probability to find E_i, V_i:

$$P\left(E_i, V_i\right) = \frac{\Omega\left(E - E_i, V - V_i\right)}{\sum_{j,k} \Omega\left(E - E_j, V - V_k\right)} = \frac{\exp\left[-\beta\left(E_i + pV_i\right)\right]}{\sum_{j,k} \exp\left[-\beta\left(E_j + pV_k\right)\right]}$$

$$\propto \exp\left[-\beta\left(E_i + pV_i\right)\right]$$

Grand-canonical ensemble: constant μ, V, T



Consider a small system that can exchange *particles* and energy with a large reservoir

$$\ln\Omega\left(N-N_{i,}E-E_i\right)=\ln\Omega\left(N,E\right)-\left(\frac{\partial\ln\Omega}{\partial E}\right)_N E_i-\left(\frac{\partial\ln\Omega}{\partial N}\right)_E N_i+\cdots$$

$$\ln\frac{\Omega\left(E-E_i,N-N_i\right)}{\Omega\left(E,N\right)}=-\frac{E_i}{k_BT}+\frac{\mu_i N_i}{k_BT}$$

Hence, the probability to find E_i,N_i:

$$P\left(E_i,N_i\right)=\frac{\Omega\left(E-E_i,N-N_i\right)}{\sum_{j,k}\Omega\left(E-E_j,N-N_k\right)}=\frac{\exp\left[-\beta\left(E_i-\mu_i N_i\right)\right]}{\sum_{j,k}\exp\left[-\beta\left(E_j-\mu_k N_k\right)\right]}$$

$$\propto\exp\left[-\beta\left(E_i-\mu_i N_i\right)\right]$$

Relating macroscopic observables to microscopic quantities

Example:

Heat capacity

Pressure

Diffusion coefficient

Computing transport coefficients from an EQUILIBRIUM simulation.

How?

Use linear response theory (i.e. study decay of fluctuations in an equilibrium system)

Linear response theory in 3 slides:

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 $\langle A \rangle_0 = \langle B \rangle_0 = 0$

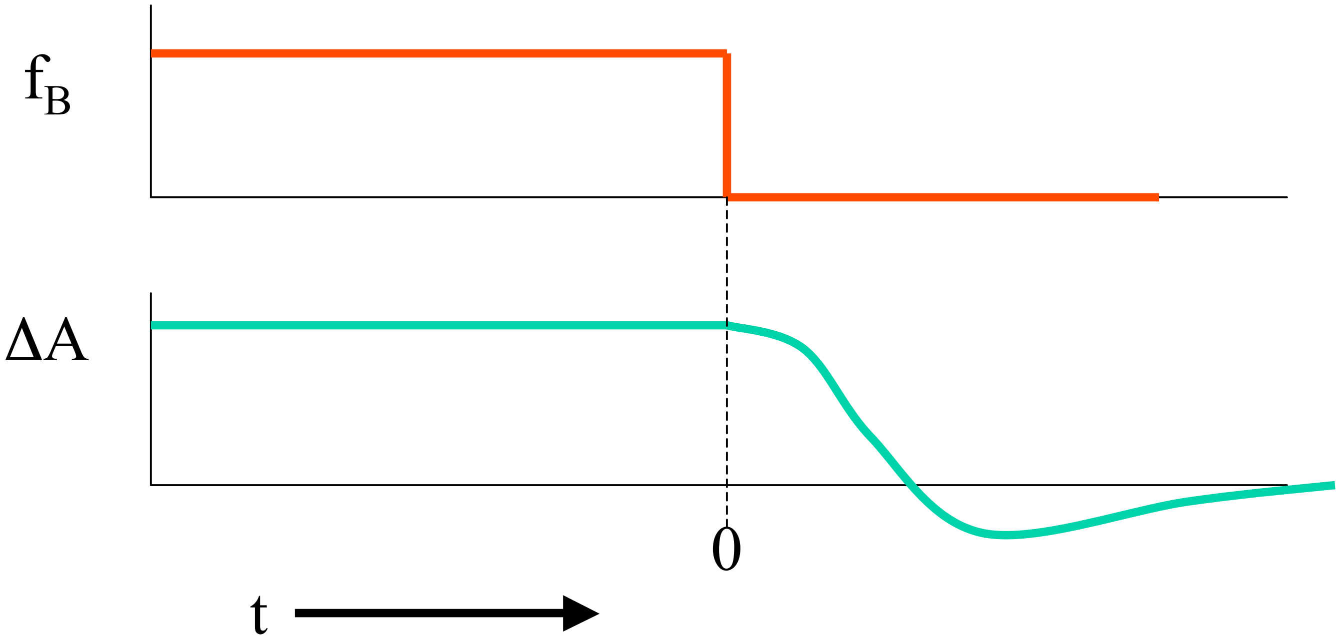
$$\langle \Delta A \rangle_{f_B} = \frac{\int \exp[-\beta(H_0 - f_B B)] A}{\int \exp[-\beta(H_0 - f_B B)]}$$

For small f_B we can linearize:

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

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:

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

Simple example: Diffusion

Average total displacement:

$$< \Delta x >= \sum_{i=1}^{N} < \Delta x_i >= 0$$

Mean squared displacement:

$$< \Delta x^2 >=< \left(\sum_{i=1}^{N} \Delta x_i \right)^2 >$$

$$= \sum_{i=1}^{N} < \Delta x_i^2 >= N a^2$$

Macroscopic diffusion equations

Fick's laws:

$$\frac{\partial c(x,t)}{\partial t} + \frac{\partial\, j_x(x,t)}{\partial x} = 0.$$

(conservation law)

$$j_x(x,t) = -D\frac{\partial\, c(x,t)}{\partial x}$$

(constitutive law)

Combine:

$$\frac{\partial\ c(x,t)}{\partial\ t} - D\frac{\partial^2\ c(x,t)}{\partial\ x^2} = 0.$$

Initial condition:

$$c(x,0) = \delta(x)$$

Solve:

$$c(x,t) = \frac{1}{(4\pi Dt)^{1/2}} \exp\left(-\frac{x^2}{4Dt}\right)$$

Compute mean-squared width:

$$\left\langle x^2(t) \right\rangle \equiv \int dx \; c(x,t) x^2$$

$$\frac{\partial}{\partial t}\int dx\; x^2 c(x,t) = D\int dx\; x^2 \frac{\partial^2 c(x,t)}{\partial\, x^2}.$$

$$\Updownarrow$$

$$\frac{d\left\langle x^2(t)\right\rangle}{dt}.$$

Integrating the left-hand side by parts:

Or:

$$2D = \lim_{t \to \infty} \frac{d \left\langle x^2(t) \right\rangle}{dt}$$

This is how Einstein proposed to measure the diffusion coefficient of Brownian particles

$$\Delta x(t) = \int_0^t dt' \; v_x(t').$$

$$2D = \lim_{t \to \infty} \frac{\partial \left\langle x^2(t) \right\rangle}{\partial t}$$

$$\left\langle x^2(t) \right\rangle = \left\langle \left(\int_0^t dt' \; v_x(t') \right)^2 \right\rangle$$

$$\left\langle \left(\int_0^t dt'\; v_x(t')\right)^2 \right\rangle =$$

$$= \int_0^t \int_0^t dt' dt'' \left\langle v_x(t') v_x(t'') \right\rangle$$

$$= 2 \int_0^t \int_0^{t'} dt' dt'' \left\langle v_x(t') v_x(t'') \right\rangle .$$

$$\left\langle v_x(t') v_x(t'') \right\rangle = \left\langle v_x(t' - t'') v_x(0) \right\rangle .$$

$$2D = \lim_{t\to\infty} 2\int_0^t dt'' \left\langle v_x(t-t'')v_x(0)\right\rangle$$

$$D = \int_0^\infty d\tau \;\; \langle v_x(\tau)v_x(0)\rangle$$

(“Green-Kubo relation”)

Other examples: shear viscosity

$$\eta = \frac{1}{V k_B T} \int_0^\infty dt \ \langle \sigma^{xy}(0) \sigma^{xy}(t) \rangle$$

$$\sigma^{xy} = \sum_{i=1}^{N} \left(m_i v_i^x v_i^y + \frac{1}{2} \sum_{j \neq i} x_{ij} f_y(r_{ij}) \right)$$

Other examples: thermal conductivity

$$\lambda_T = \frac{1}{V k_B T^2} \int_0^\infty dt \ \langle j_z^e(0) j_z^e(t) \rangle$$

$$j_z^e = \frac{d}{dt} \sum_{i=1}^{N} z_i \frac{1}{2} \left(m_i v_i^2 + \sum_{j \neq i} v(r_{ij}) \right)$$

Other examples: electrical conductivity

$$\sigma_e = \frac{1}{V k_B T} \int_0^\infty dt \ \left\langle j_x^{\mathrm{el}}(0) j_x^{\mathrm{el}}(t) \right\rangle$$

$$j_x^{\mathrm{el}} = \sum_{i=1}^{N} q_i v_i^x .$$

The Monte Carlo Method

Aim: to compute thermal averages of equilibrium systems.

$$\langle A \rangle = \frac{\sum_i \exp(-\epsilon_i/k_B T) A_i}{\sum_i \exp(-\epsilon_i/k_B T)}$$

Where *i* labels all eigenstates of the system, and

$$A_i = \langle i | A | i \rangle$$

Classical limit: replace the SUM over quantum states by an INTEGRAL of phase space

$$\langle A \rangle = \frac{\int \mathrm{d}\mathbf{p}^N \mathrm{d}\mathbf{r}^N \; A(\mathbf{p}^N, \mathbf{r}^N) \exp[-\beta \mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)]}{\int \mathrm{d}\mathbf{p}^N \mathrm{d}\mathbf{r}^N \; \exp[-\beta \mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)]}$$

Where H is the Hamiltonian of the system and $\beta=1/kT$

In replacing the sum by an integral, we have attributed a "volume" h^{3N} to every quantum state

Problem:

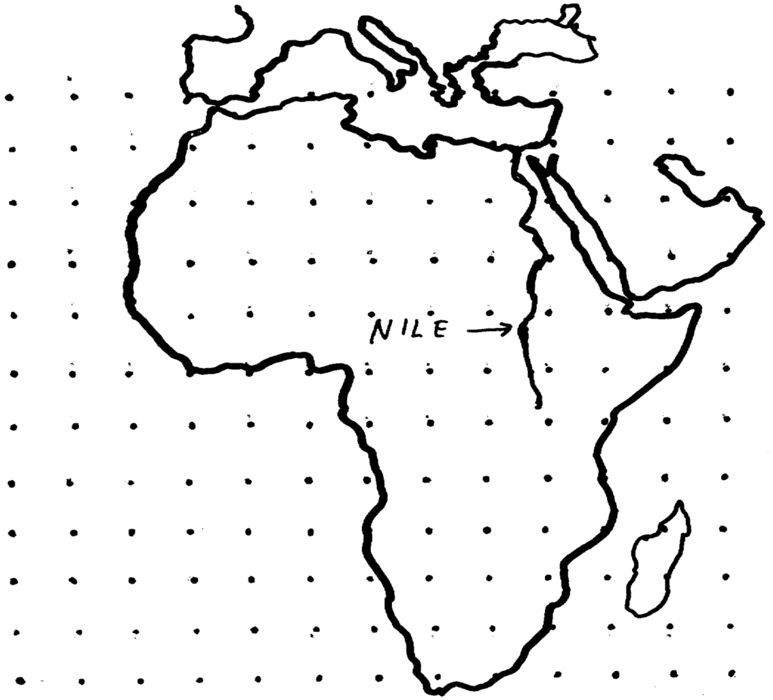
We cannot compute the sum over all quantum states (because there are so many)

And we cannot compute the classical integral either (except the integration over momenta).

Consider "normal" numerical integration

100 particles, 3 dimensions, 10 points in every direction.

Requires 10^{300} points for a very poor estimate of the integral…



Similar problem (but much less serious):

Measure the depth of the Nile by quadrature…

Brute-force Monte Carlo (“Random sampling”)

$$\int_V dx_1 \cdots dx_N f(x) \approx V^N \sum_{i-1}^{M} f(r_i)$$

(M random points r_i in (hyper)volume V_N)

BETTER
STRATEGY:

IMPORTANCE

SAMPLING

We wish to perform a RANDOM WALK in configuration space, such that

The number of times that each point is visited, is proportional to its Boltzmann weight.

$$n(\mathbf{r}^N) = c\,\mathsf{exp}[-\beta\mathcal{U}(\mathbf{r}^N)]$$

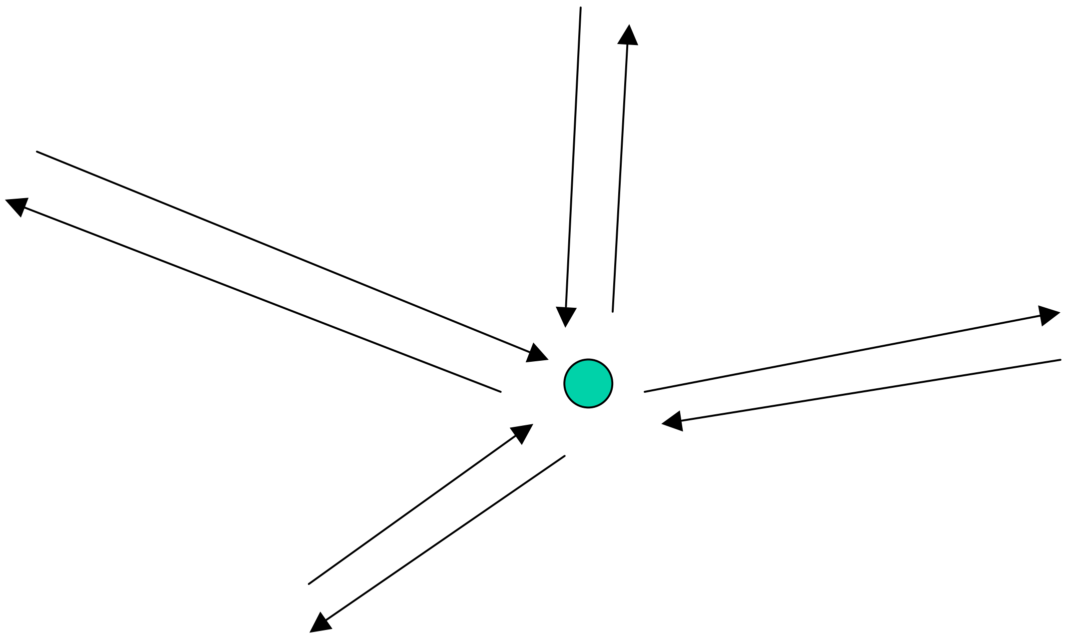
Then

$$\langle A \rangle \approx \frac{1}{L} \sum_{i=1}^{L} n_i \, A(\mathbf{r}_i^N).$$

How do we achieve that?

Whatever our rule is for moving from one point to another, it should not destroy the equilibrium distribution.

That is: in equilibrium we must have



$$\mathcal{N}(o)\sum_n \pi(o \to n) = \sum_n \mathcal{N}(n)\pi(n \to o)$$

Stronger condition:

$$\mathcal{N}(o)\pi(o \to n) = \mathcal{N}(n)\pi(n \to o).$$

For every pair $\{n,o\}$.

Detailed Balance

Now we construct the transition probabilities

$$\pi(o \to n) = \alpha(o \to n) \times \mathsf{acc}(o \to n)$$

Then, detailed balance implies that:

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

$$=$$

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

Often, we choose

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

Then it follows that

$$\mathcal{N}(o) \times \mathsf{acc}(o \rightarrow n)$$

$$=$$

$$\mathcal{N}(n) \times \mathsf{acc}(n \rightarrow o)$$

$$\frac{\mathsf{acc}(o \rightarrow n)}{\mathsf{acc}(n \rightarrow o)} = \frac{\mathcal{N}(n)}{\mathcal{N}(o)} = \mathsf{exp}\{-\beta[\mathcal{U}(n) - \mathcal{U}(o)]\}$$

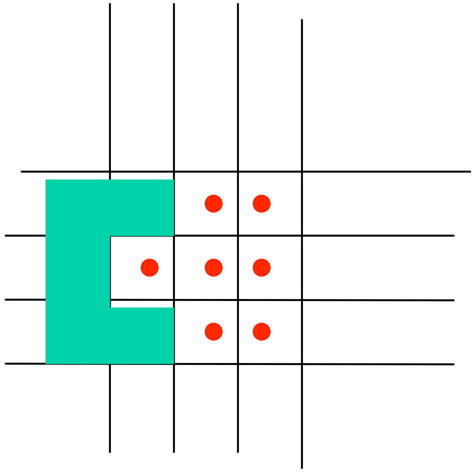
Metropolis, Rosenbluth,Rosenbluth,

Teller and Teller choice:

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

Kirkwood’s objection:

“If a trial move has been rejected, one should not count the original state AGAIN…”



Counter-example:

Ideal gas on a lattice.

Problem for both MC and MD:

“non-ergodicity”

(i.e. the sampling is limited to a subset of all possible states of the system)

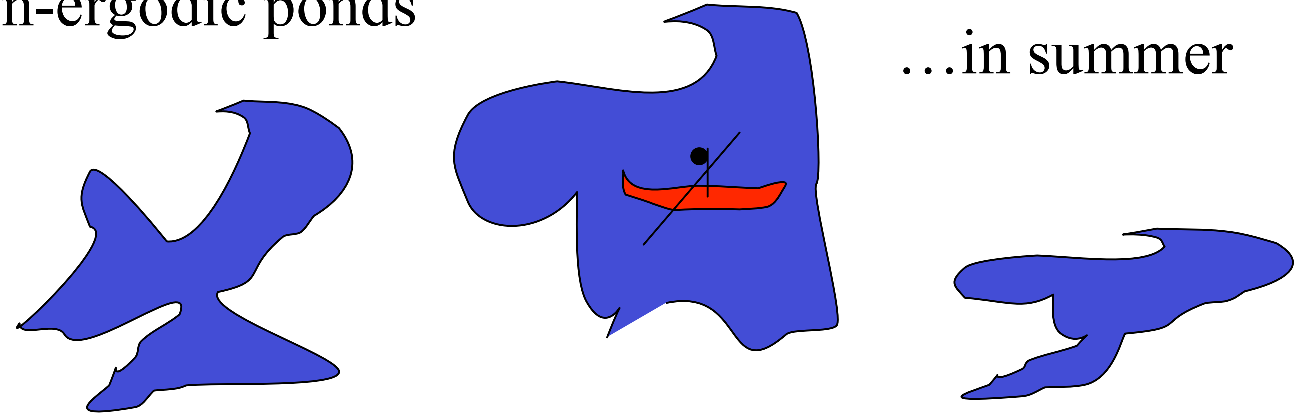
Often, ergodicity problems occur at low temperatures

(glasses, gels, disordered crystals, …)

ANALOGY:

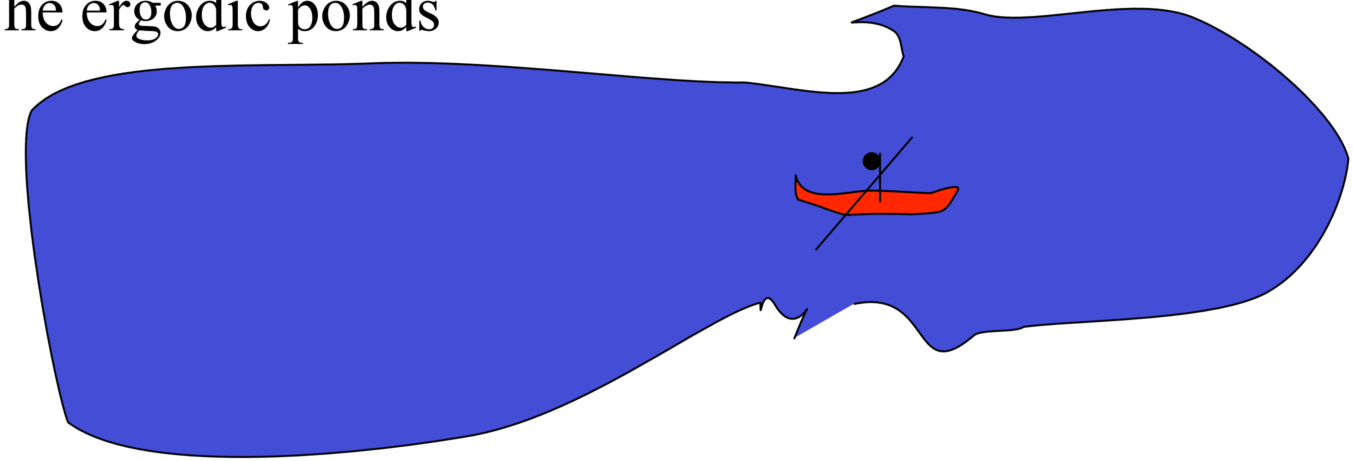
The Non-ergodic ponds

…in summer



…in spring

The ergodic ponds



In Statistical Mechanics language:

“Glassy” energy landscapes.

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

At high temperatures: one “ergodic” lake.

Parallel Tempering

COMBINE “summer”and “spring” 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

$$\frac{\text{new Boltzmann weight}}{\text{old Boltzmann weight}}$$

$$=$$

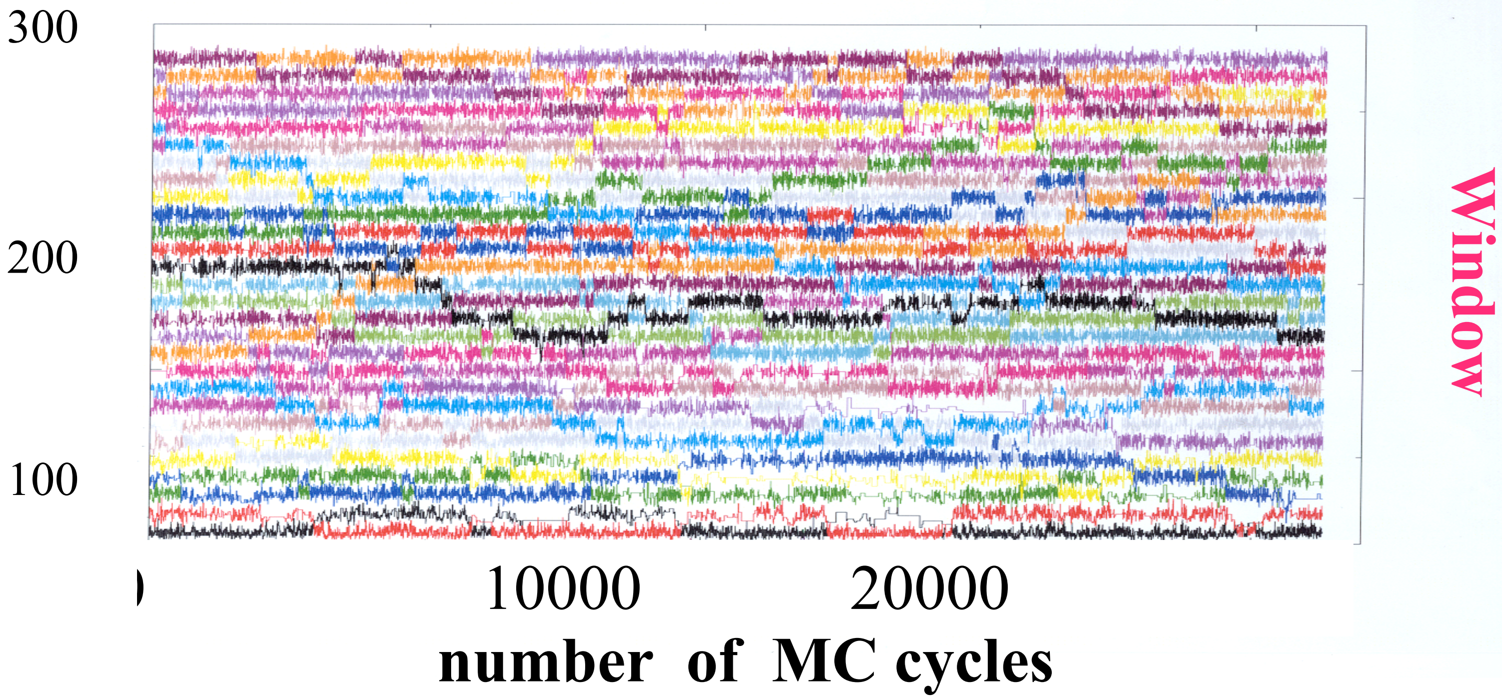
$$\exp\{-(\beta_1-\beta_2)[\mathcal{U}_2(\mathbf{r}^N)-\mathcal{U}_1(\mathbf{r}^N)]$$

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

$$\mathsf{acc(swap)}$$

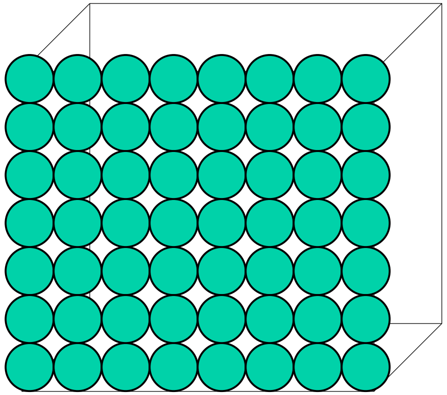
$$=$$

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


300
200
100
10000
20000
number of MC cycles
Window

Practical issues:

1. Boundary conditions
2. Time-saving devices

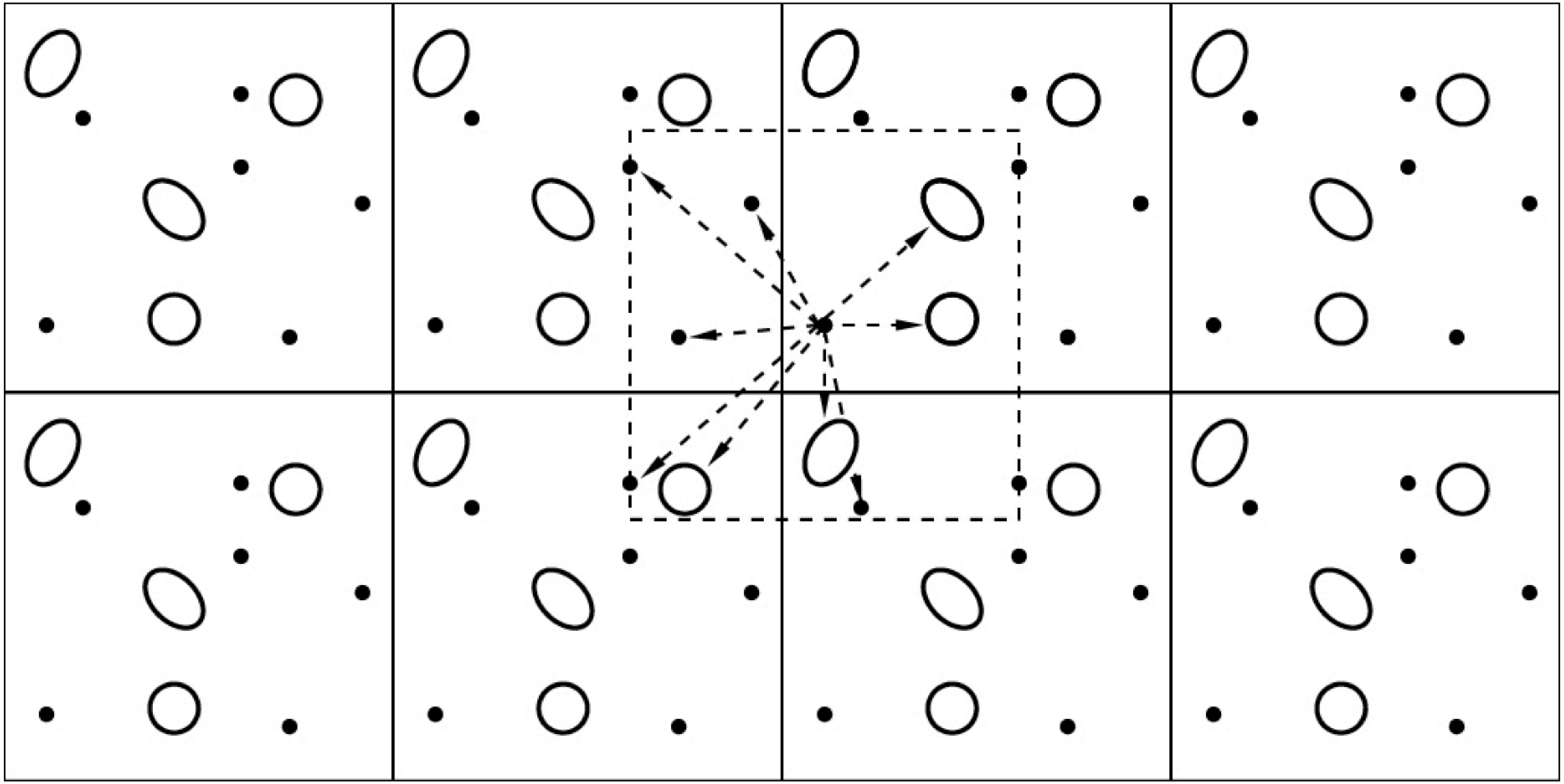


In small systems, boundary effects are always large.

1000 atoms in a simple cubic crystal – 488 boundary atoms.

1000000 atoms in a simple cubic crystal – still 6% boundary atoms…

“Solution” : Periodic boundary conditions



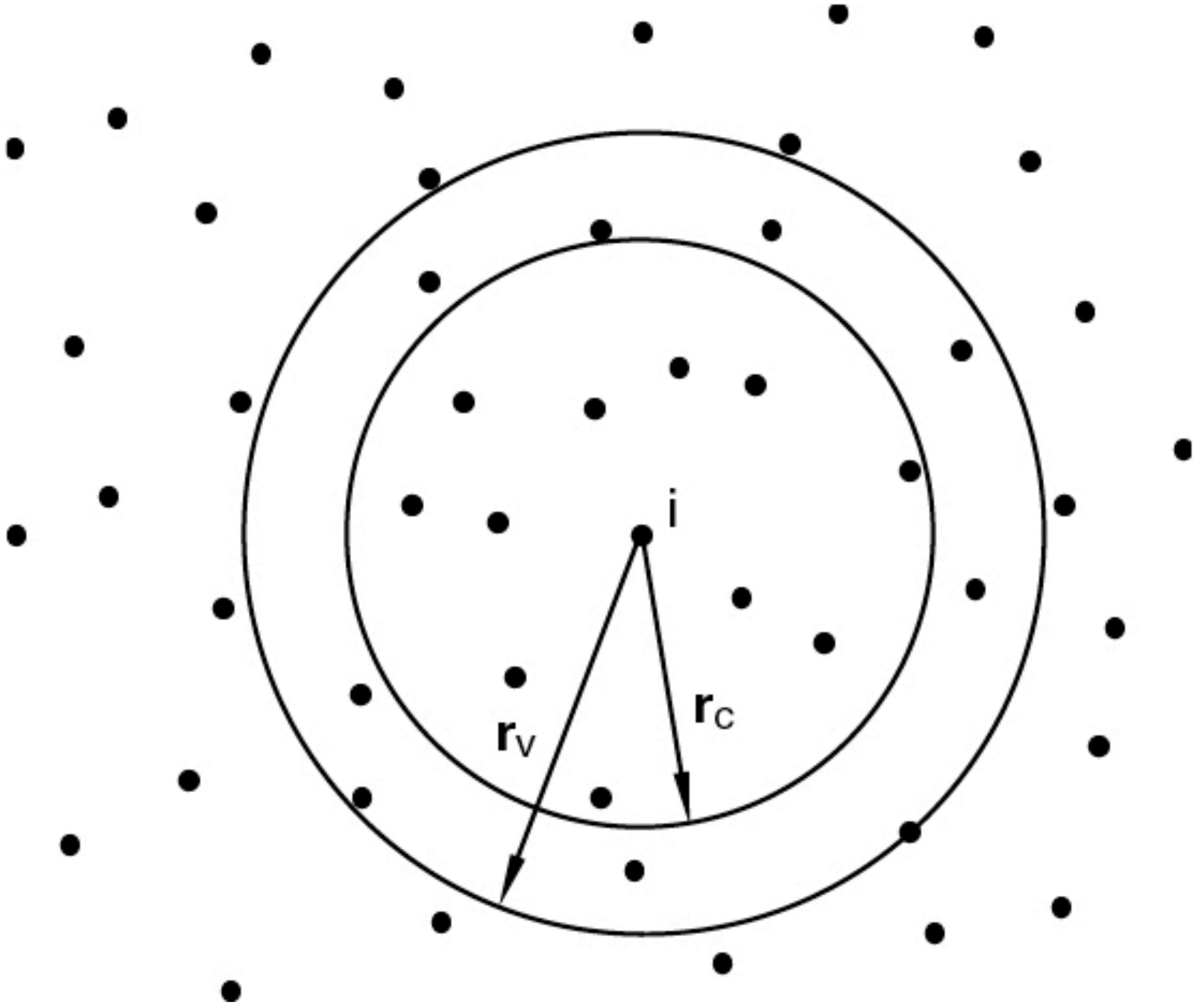
The most time-consuming part of any simulation is the evaulation of all the interactions between the molecules.

In general: $N(N-1)/2 = O(N^2)$

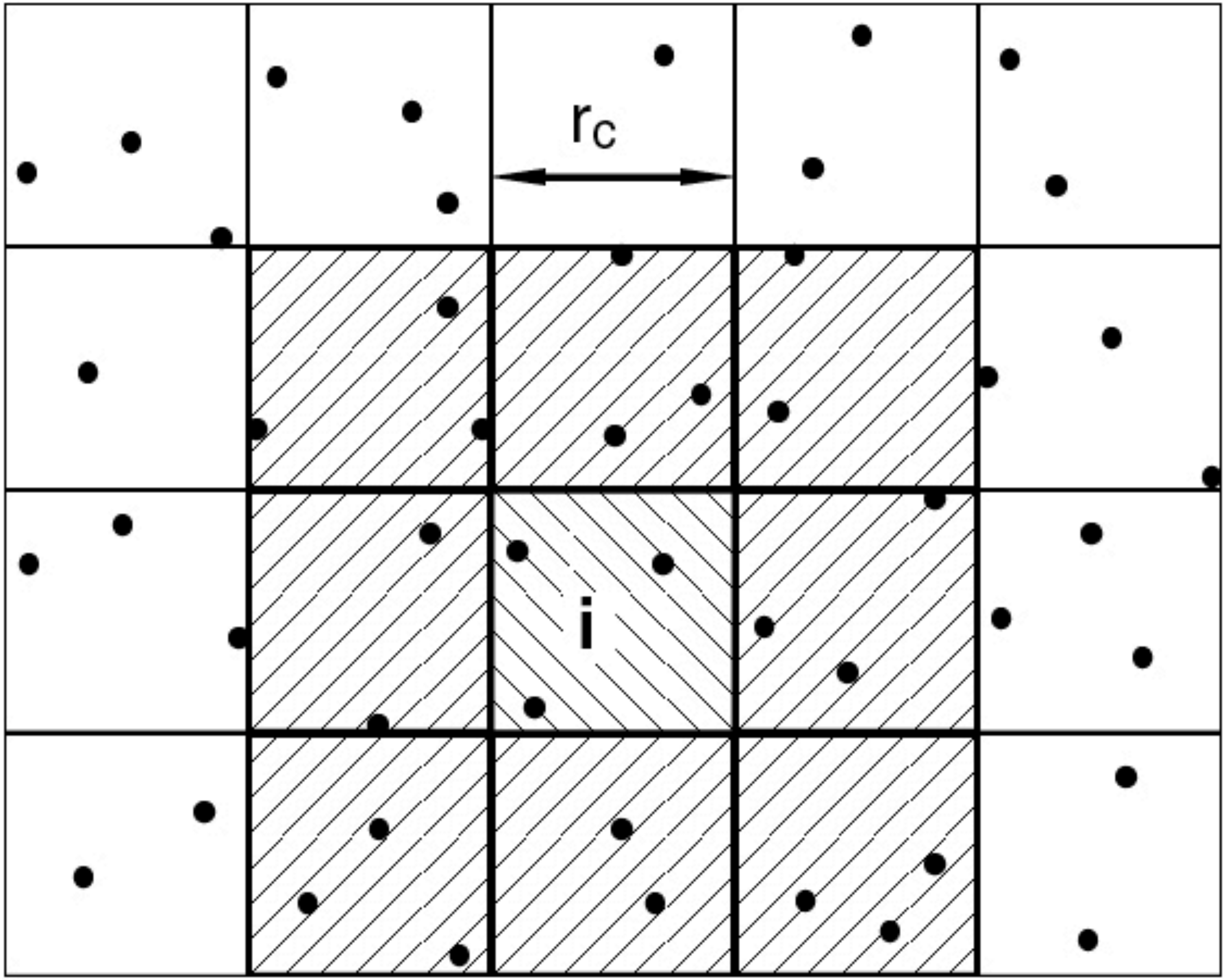
But often, intermolecular forces have a short range:

Therefore, we do not have to consider interactions with far-away atoms…

Verlet list



Link list



NOTE:

Long-ranged forces require special techniques.

1. Coulomb interaction (1/r in 3D)
2. Dipolar interaction ($1/r^3$ in 3D)

…and, in a different context:

1. Interactions through elastic stresses (1/r in 3D)
2. Hydrodynamic interactions (1/r in 3D)
3. …

Beyond standard Monte Carlo

Metropolis Monte Carlo:

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

Metropolis, Rosenbluth,Rosenbluth, Teller and Teller:

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

Alternative: “symmetric rule”

$$\mathrm{acc}(o \rightarrow n) = \frac{\exp\{-\beta\mathcal{U}(\mathbf{r}'^N)\}}{\exp\{-\beta\mathcal{U}(\mathbf{r}'^N)\} + \exp\{-\beta\mathcal{U}(\mathbf{r}^N)\}}$$

Unsatisfactory?


MAXWELL DEMONS ON STRIKE
100% NO LESS!
ACCEPT OUR MOVES NOW!!

Solution of conflict: if we do **not** impose

$$\alpha(o \rightarrow n) = \alpha(n \rightarrow o)$$ then

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

In particular, if:

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

Then

$$\frac{\mathrm{acc}(o \rightarrow n)}{\mathrm{acc}(n \rightarrow o)} = 1$$ (100% acceptance)

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

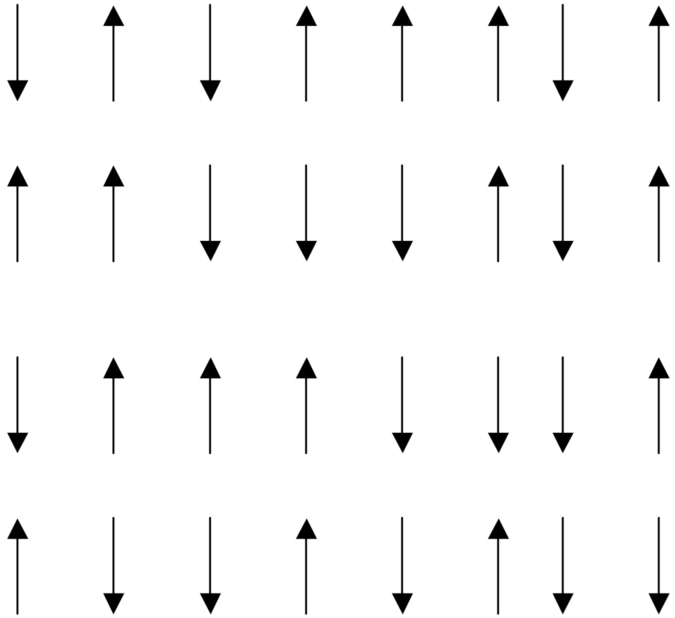
Discrete spin models (Potts, Ising).

Illustration: 2D Ising model:

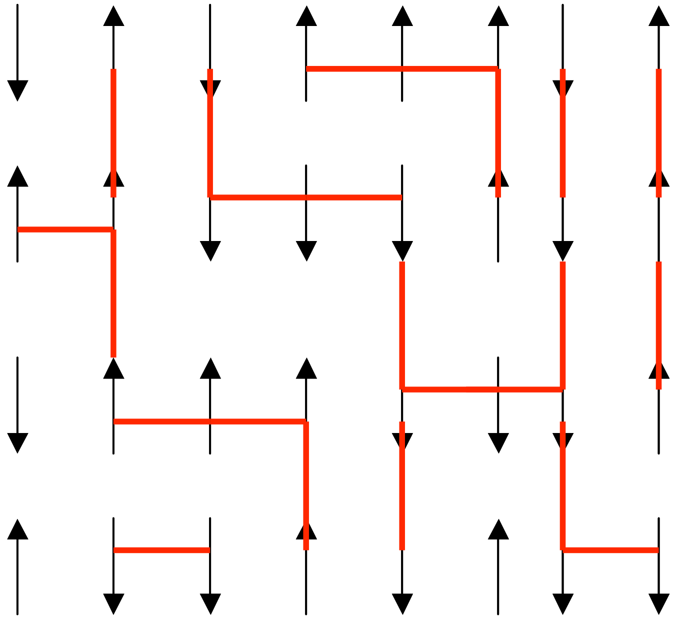
Parallel nearest neighbor spins: energy **–J**

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

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



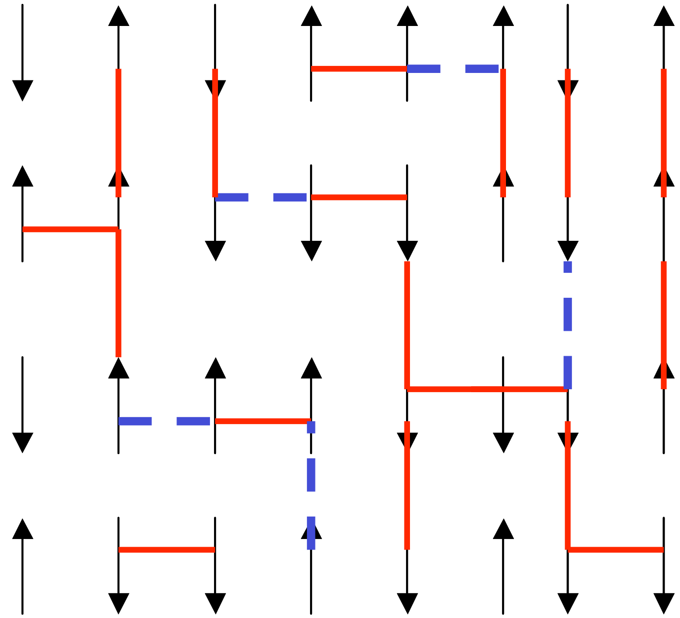
Snapshot: some neighbors are parallel, others anti-parallel



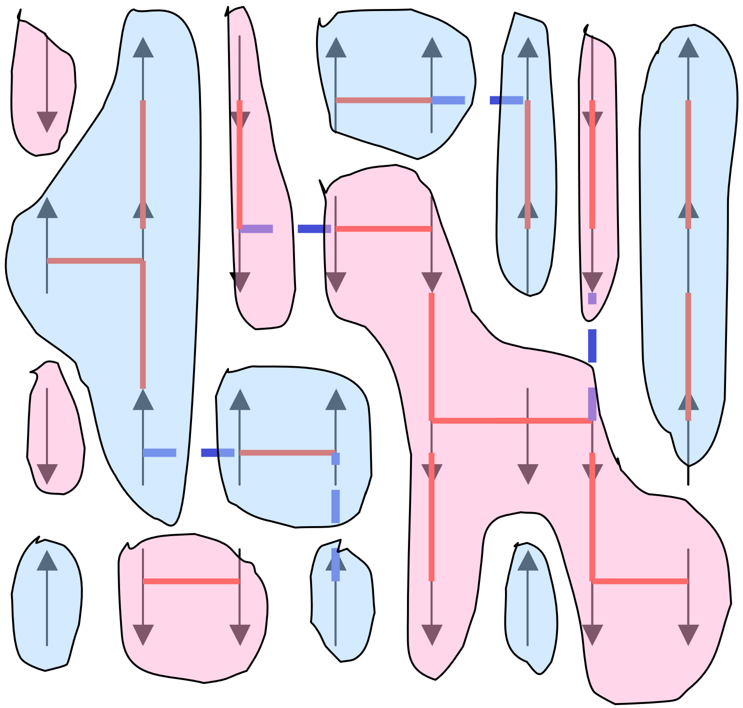
Count number of bonds between parallel neighbors: **N_p**

Number of bonds between anti-parallel neigbors is: **N_a**

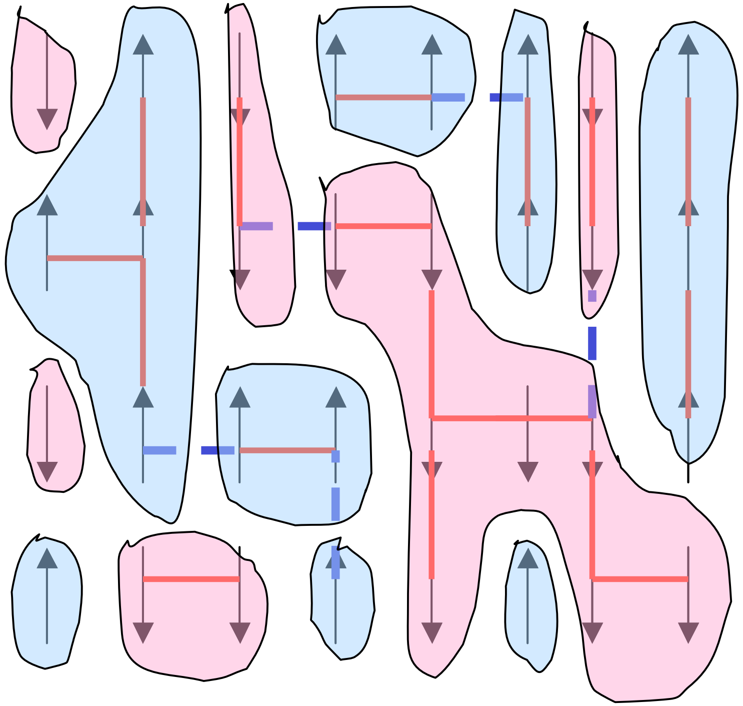
Total energy: **$U = (N_a - N_p)\ J$**



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).

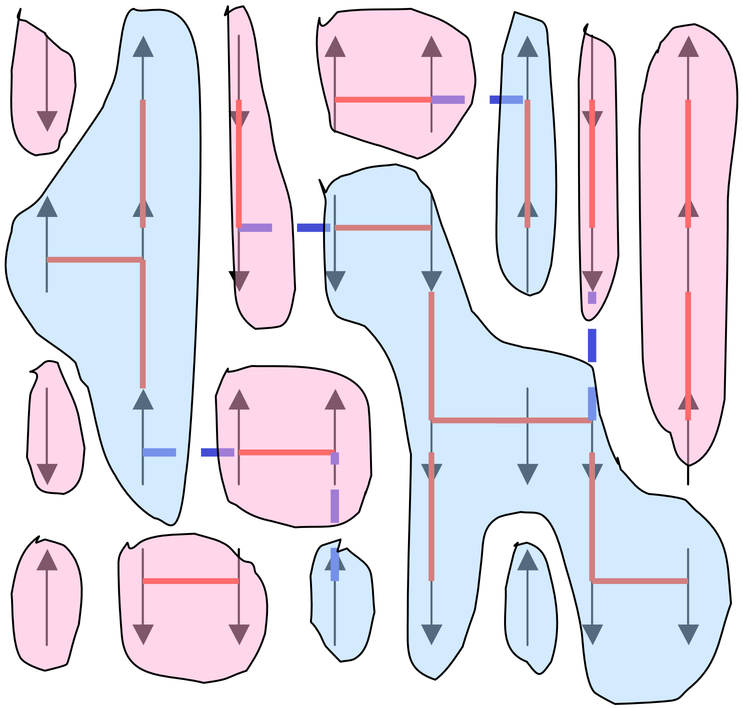


Form clusters of all spins that are connected by bonds. Some clusters are all “spin up” others are all “spin down”. Let us denote the number of clusters by **M**.



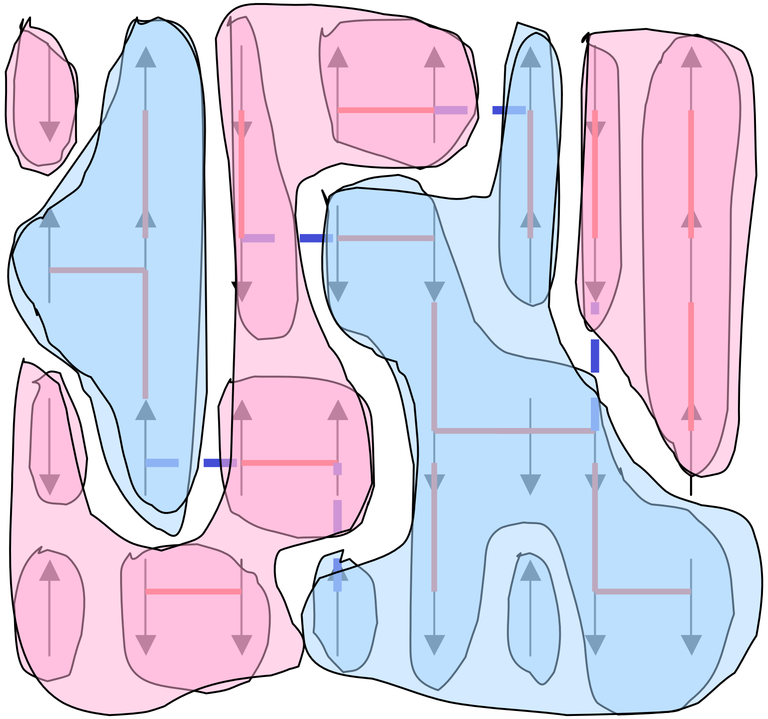
The probability to generate a particular cluster structure where there are $\mathbf{n_c}$ bonds between $\mathbf{N_p}$ pairs of parallel neighbors is:

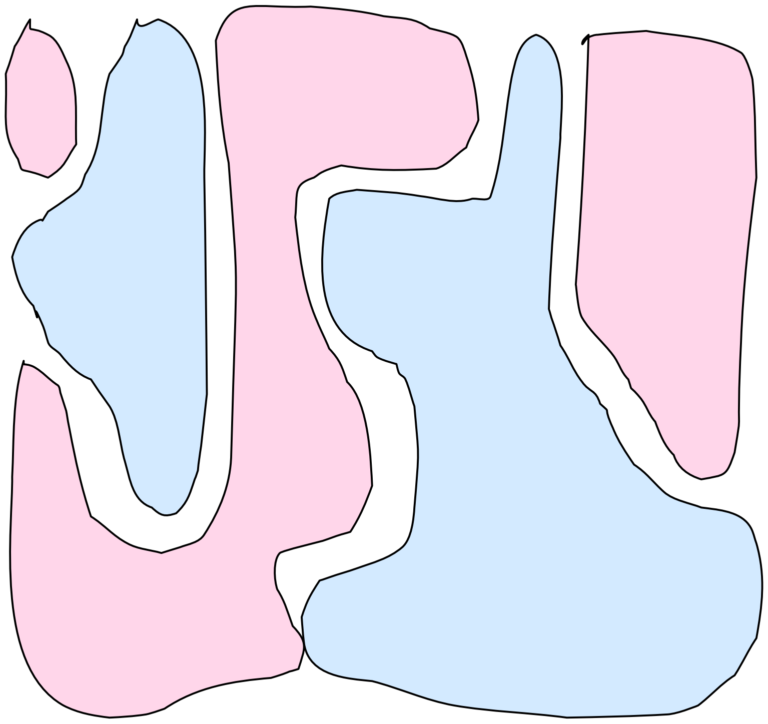
$$P_{gen} = p^{n_c}(1-p)^{N_p - n_c}$$



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:

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

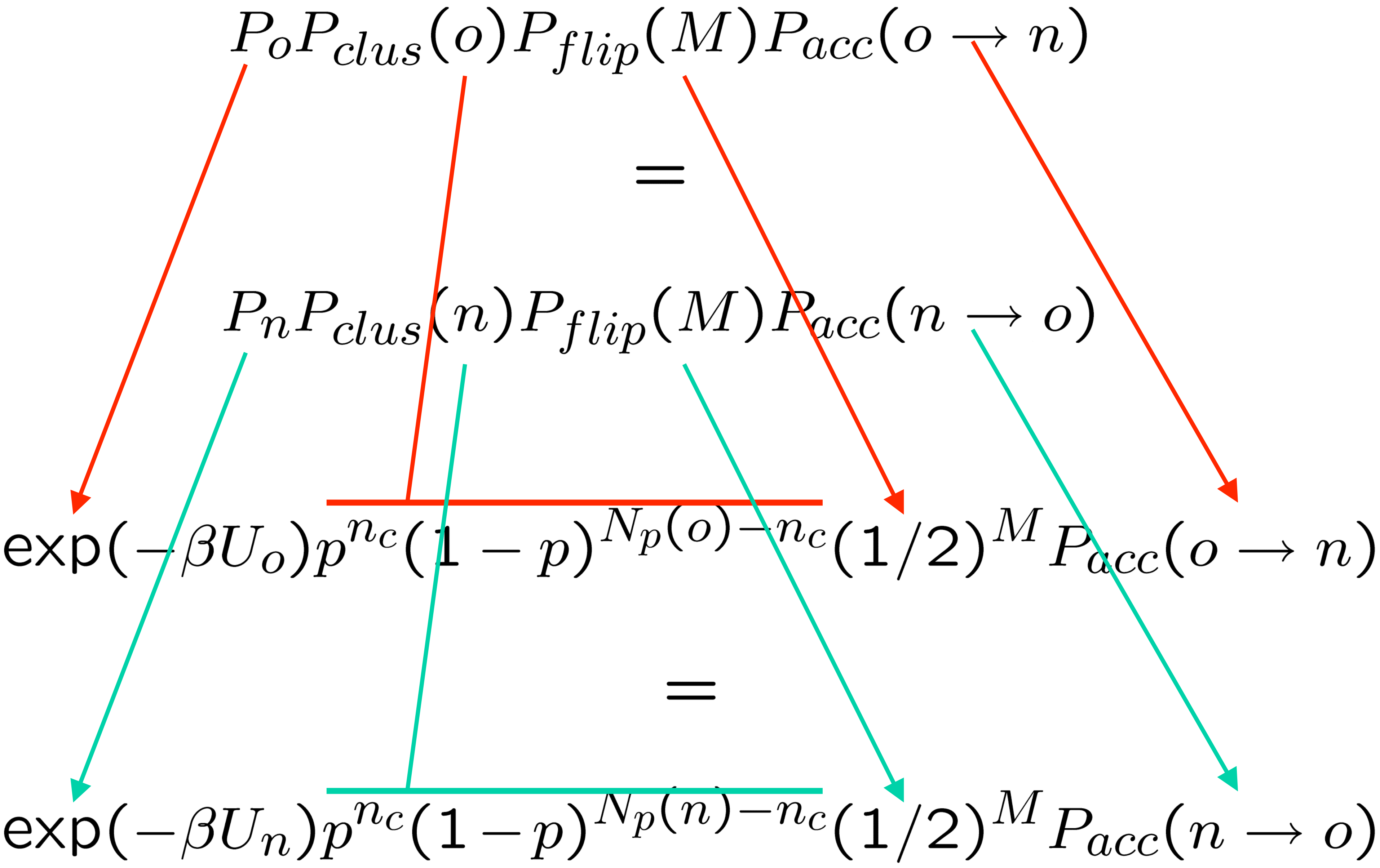
$$=$$

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

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

$$=$$

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



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

$$=$$

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

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

$$=$$

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

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

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

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

$$=$$

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

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

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

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

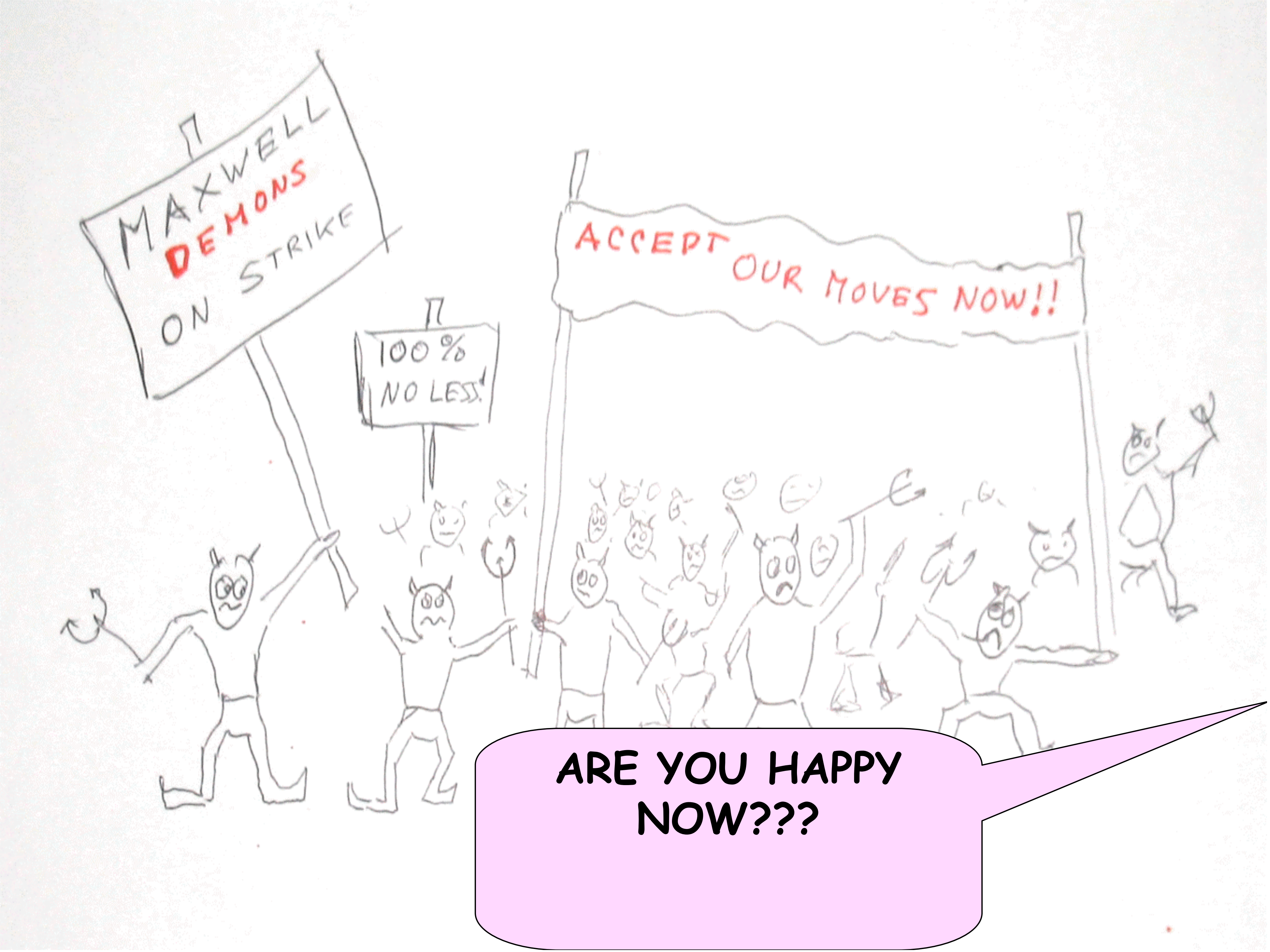
Combining this with:

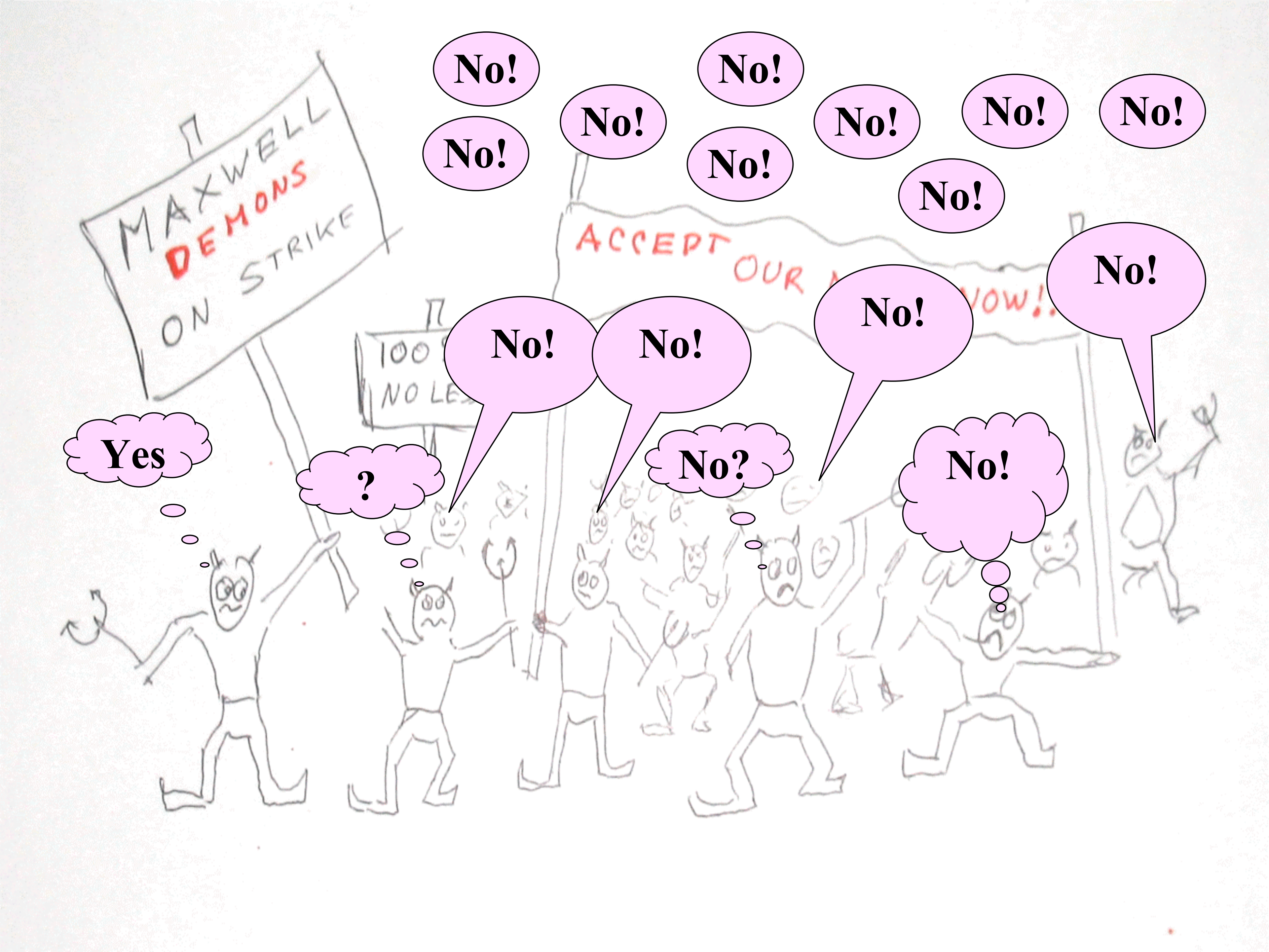
$$\exp(\beta(U_n - U_o)) = (1-p)^{N_p(n)-N_p(o)}$$

we obtain:

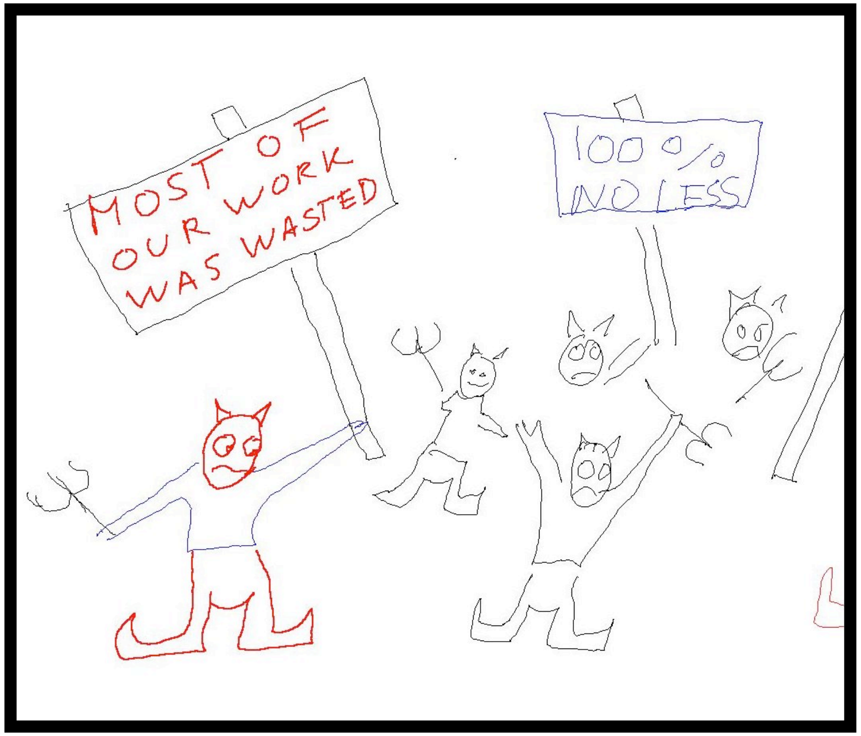
$$p = 1 - \exp(-2\beta J)$$

100% acceptance!!!


MAXWELL DEMONS ON STRIKE
100% NO LESS!
ACCEPT OUR MOVES NOW!!
ARE YOU HAPPY NOW???


MAXWELL DEMONS ON STRIKE
100
NO LE
ACCEPT OUR
NOW!!
No!
No!
No!
No!
No!
No!
No!
No!
No!
No!
No!
No!
No!
Yes
?
No?
No!

Why not?



If we construct **n** clusters and we can flip each one independently, then we have generated **2^n** possible states…

…and yet we accept only **one**!

Waste Recycling in MC SAMPLING

Include "rejected" moves in the sampling

Dangerous?:

Metropolis "importance" sampling is **based** on the earlier (Ulam/von Neumann) **rejection** method applied to **random** MC sampling

This is the key:

$$\sum_m \rho(m)\pi_{mn} = \rho(n)$$

The transition matrix π leaves the equilibrium distribution ρ unchanged.

$$\langle A \rangle_{\rho} = \sum_{n} A_n \rho_n$$

This, we can rewrite as:

$$\sum_{n} A_n \rho_n = \sum_{n} \sum_{m} A_n \rho_m \pi_{mn} = \sum_{m} \rho_m \sum_{n} A_n \pi_{mn}$$

$$= \sum_{m} \rho_m \sum_{n} A_n \pi_{mn} \Leftrightarrow \langle A \rangle_{\rho} = \left\langle \sum_{n} \pi_{mn} A_n \right\rangle_{\rho_m}$$