Why Ω works for large N

Derivation of the Boltzmann Distribution Law

Partition Function

• Why Ω works for large N

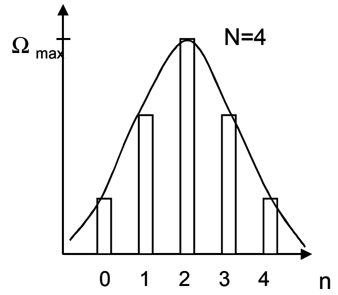
We have seen that a system will vary its degrees of freedom in order to maximize Ω and thus S. A system has a higher probability of being in a state due to it being more probable. This allows us to simply count states and see which one is more likely.

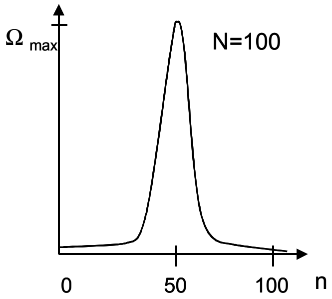
The lattice model of mixing gases had only N=8 particles. Is this approach still justified when we look at a larger number of particles, like N_A? It turns out the most probable state at low N becomes even more likely at very high N.

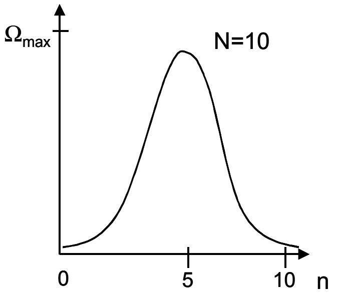
Consider: coin flips

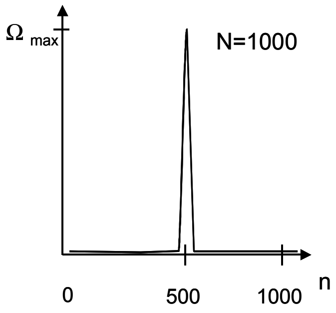
n_H	Ω	S =klnΩ
4	$\Omega = \frac{N!}{n!(N-n)!} = \frac{4!}{4!0!} = 1$	0
3	$\Omega == \frac{4!}{3!1!} = 4$	1.386k
2	$\Omega == \frac{4!}{2!2!} = 6$	1.792k
1	$\Omega == \frac{4!}{1!3!} = 4$	1.386k
0	$\Omega == \frac{4!}{0!4!} = 1$	0

Then do for N = 10, 100, 1000









Ω becomes increasingly narrower as N↑. Compare numerically:

$$\Omega(5,10) = \frac{10!}{5!5!} = 252$$

5.6X more likely

$$\Omega(2,10) = \frac{10!}{2!8!} = 45$$

$$\Omega(50,100) = \frac{100!}{50!50!} = 1\times 10^{29}$$

10^9X more likely!

$$\Omega(20,100) = \frac{100!}{20!80!} = 5\times 10^{20}$$

Even though the process is totally random: If the number of trials N is large enough, the composition of the outcomes becomes predictable with great precision.

This allows us to better predict the most probable state!

maximizing Ω = maximizing S

• Derivation of the Boltzmann Distribution Law

Microscopic definition of entropy:

$$S = -k\sum_{i=1}^{t} p_j \ln p_j$$

What probability distribution (set of p_j's) maximizes S?

$$\frac{\partial S}{\partial p_j} = 0 \quad \text{for all j}$$

constraint that probabilities sum to 1:

$$\sum_{j=1} p_j = 1$$

$$\sum_{j=1}^{t} dp_j = 0$$

Utilize **Lagrange multipliers** to solve this problem. We add the constraint to the equation we are trying to maximize with a multiplier, α. Then when we maximize the resulting equation the value of α is determined. i.e., solving the set of equations:

$$\sum_{j=1}^{t}\left[\left(\frac{\partial S}{\partial p_j}\right) - \alpha\right] dp_j = 0 \quad \text{for all j}$$

Plug in definition of S (pj):

$$\frac{\partial}{\partial p_j}\left(-k\sum_{j}^{t} p_j \ln p_j - \alpha\left(\sum_{j=1}^{t} p_j\right)\right) = 0$$

Take the derivative:

$$-k(\ln p_j + 1) - \alpha = 0$$

$$\ln p_j = -\frac{\alpha}{k} - 1$$

$$p_j = e^{\left(\frac{-\alpha}{k}-1\right)}$$

Divide pi by 1 to get rid of α:

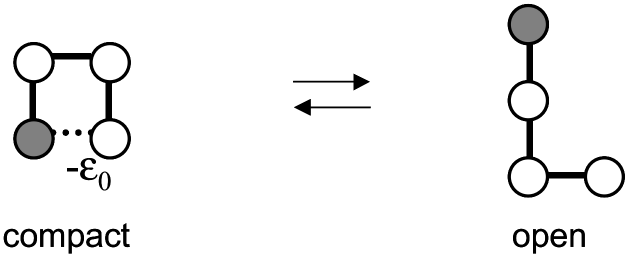
$$\frac{p_j}{\sum_{i=1}^{t} p_i} = \frac{e^{\left(\frac{-\alpha}{k}-1\right)}}{te^{\left(\frac{-\alpha}{k}-1\right)}} = \frac{1}{t}$$

This says: **Flattest probability distributions have highest S**. This is something we already knew.

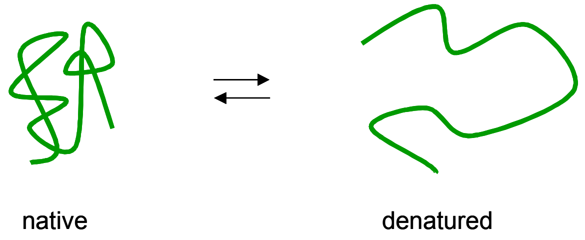
Now, what happens when we impose a constraint on the system? i.e., you have a given temperature, and can sum up to a particular total energy. This is a more realistic problem to solve.

Let's put this into practice with an example.

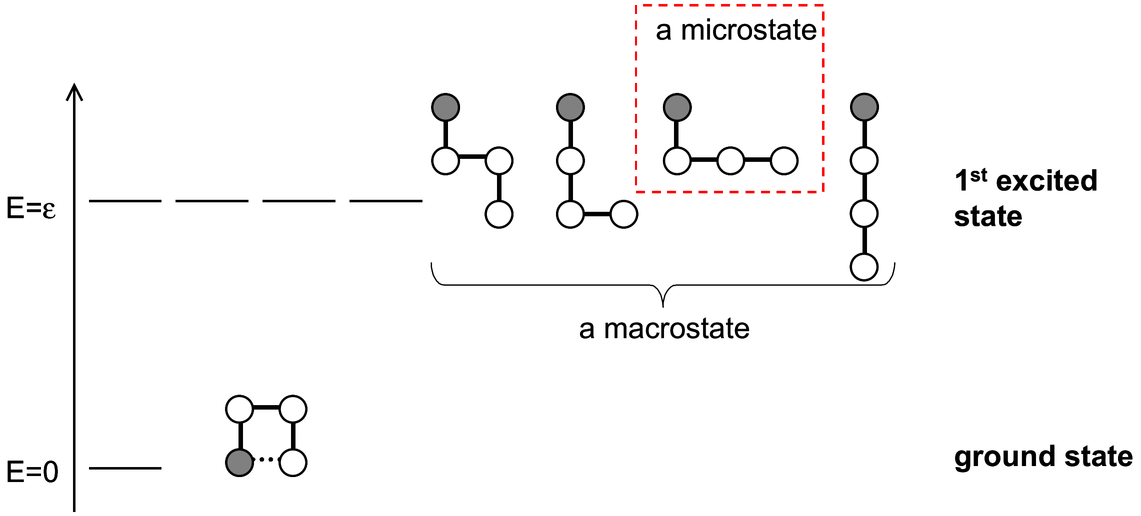
Simple model to illustrate: **4 bead polymer**



The polymer can assume multiple configurations. We'll label one end atom so that it is distinguishable from the other atoms in the chain. The polymer is stabilized when in compact configuration by energy ε from open state. This is represented by the dashed line. This is a simple model utilized by those studying protein folding as it can represent the configurations of a protein in the folded and unfolded states. It represents a polypeptide chain that has only 4 amino acids, and a great simplification of real proteins in that the chain can assume only a small number of conformations: one compact and four open.



One end bead is labeled so that it is distinguishable from other end.



microstate: a possible configuration. “snapshot.” A measurement averages over several microstates

macrostate: a collection of microstates with the same energy

Define the E=0 state as the compact form, where the chain is stabilized by some energy ε relative to the open state due to the interaction between bead 1 and bead 4.

Degree of freedom: physical conformation of the chain and the energy of each conformation, or microstate.

What is the probability distribution that minimizes or maximizes a relevant thermodynamic quantity? What happens if we try do this in real laboratory conditions? (T,V,N) or (T,P,N) controlled.

Let’s say we have (T,V,N) constant, making A what we want to minimize.

$$dA = dU = TdS$$

at equilibrium:

$$dA = 0$$

Goal: Get dU and dS and solve for p_j that makes dA = 0.

Differentiate with respect to p_i

$$dS = -k\sum_{j=1}^{t}\left(1+\ln p_j\right)dp_j$$

Recall definition of an average value:

$$\langle E\rangle = \sum_{j=1}^{t} p_j E_j$$

$$dU = d\langle E\rangle = \sum_{j=1}^{t}\left(p_j dE_j + E_j dp_j\right)$$

Energy levels do not depend on T. p_j, or how they are populated, do.

$$dU = d\langle E\rangle = \sum_{j=1}^{t}\left(E_j dp_j\right)$$

$$dA = d\langle E\rangle - TdS = 0$$

use the constraint:

$$\sum_{j=1}^{t} p_j = 1$$

Which allows us to use the Lagrange Multiplier

$$\alpha\sum_{j=1}^{t} dp_j = 0$$

Plug everything back into dA equation:

$$dA = d\langle E\rangle - TdS = 0$$

$$= \sum_{j=1}^{t} E_j dp_j - T\left(-k\sum_{j=1}^{t}\left(1+\ln p_j\right)dp_j\right) + \underbrace{\alpha\sum_{j=1}^{t} dp_j}_{\text{adding 0}} = 0$$

group dp_j terms:

$$dA = \sum_{j=1}^{t}\underbrace{\left[E_j + kT\left(1+\ln p_j\right)+\alpha\right]}_{\text{must }=0}dp_j = 0$$

$$\ln p_j^* = -\frac{E_j}{kT} - \frac{\alpha}{kT} - 1$$

p_j*= set of p_j that satisfies dA=0

$$p_j^* = \exp\left(-\frac{E_j}{kT}\right)\exp\left(-\frac{\alpha}{kT} - 1\right)$$

We'll eliminate α from the equation by using

$$\sum_{j=1}^{t} p_j = 1$$

First, sum both sides:

$$\sum_{j=1}^{t} p_j^* = 1 = \sum_{j=1}^{t} \exp\left(-\frac{E_j}{kT}\right)\exp\left(-\frac{\alpha}{kT} - 1\right)$$

$$1 = \exp\left(-\frac{\alpha}{kT} - 1\right)\sum_{j=1}^{t} \exp\left(-\frac{E_j}{kT}\right)$$

Rearranging the last expression

$$\frac{1}{\sum_{j=1}^{t} \exp\left(-\frac{E_j}{kT}\right)} = \exp\left(-\frac{\alpha}{kT} - 1\right)$$

Plug this back into p_j*:

$$\boxed{\begin{aligned} p_j^* &= \frac{\exp\left(-\frac{E_j}{kT}\right)}{\sum_{j}^{t} \exp\left(-\frac{E_j}{kT}\right)} \\ &= \frac{\exp\left(-\frac{E_j}{kT}\right)}{Q} \end{aligned}}$$

Boltzmann Distribution Law

p_j is the probability that the systems is in the E_j^{th} energy level

We have defined the denominator as **Q, the partition function**

$$\boxed{\frac{1}{Q} \equiv \frac{1}{\sum_{j=1}^{t} \exp\left(-\frac{E_j}{kT}\right)}}$$

We arrived here by finding the probability distribution, or set of p_j's, that minimizes the free energy.

What does it say?

- When you are trying to maximize entropy, minimize energy: more particles like to have lower energies. Particles populate relatively low E_j apiece
- Probability distributions have an exponential form when you place constraints on them (not flat, like for the case of no constraints)

Relative populations of two levels:

$$\frac{p_i^*}{p_j^*} = \exp\left(-\left(E_i - E_j\right) \Big/ kT\right)$$

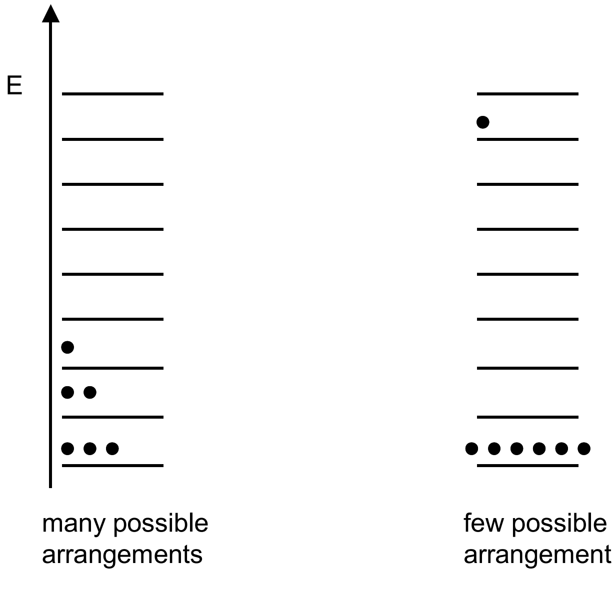
If j higher than i, then $E_i - E_j < 0$ (negative)

so exp (+)

$p_i/p_j > 1$, ∴ more in ith level.

Note: Particles do not have a preference for the lower energies, there is just a greater number of ways to arrange the particles so that they distribute the E.

For a given E_{tot}, can arrange particles in several ways to achieve E_{tot}. However, the Boltzmann Distribution Law says that the left hand situation is much more probable as it has the higher entropy.



• What is the partition function?

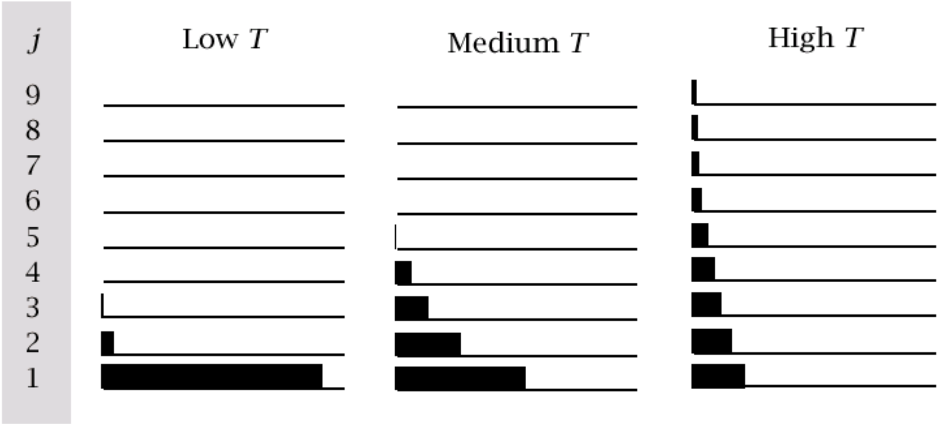
In our derivation of the Boltzmann equation, the partition function, Q, came out.

$$Q \equiv \sum_{j=1}^{t} \exp\left(-\frac{E_j}{kT}\right)$$

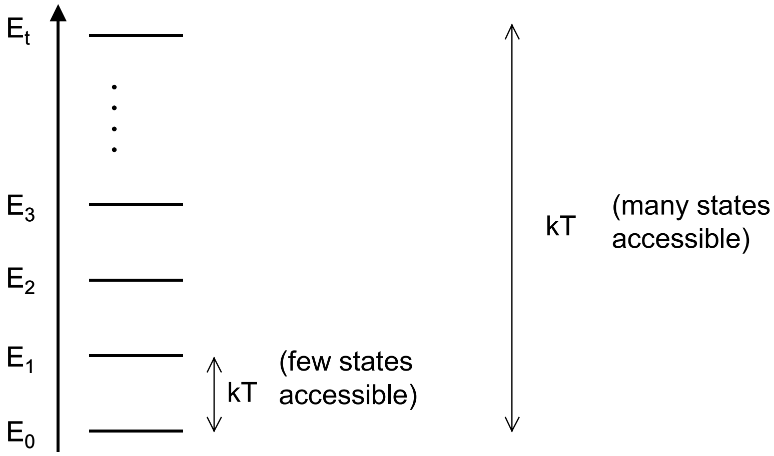
Q describes how the particles are partitioned throughout accessible states. It is a number. Note that Q is temperature dependent!

In simpler terms: Q tells you the number of states that are **effectively** accessible to the system at a given temperature.

Qualitatively:



If you have t energy levels:



$$Q \equiv \sum_{j=1}^{t} \exp\left(-\frac{E_j}{kT}\right) = \exp\left(-\frac{E_0}{kT}\right) + \exp\left(-\frac{E_1}{kT}\right) + \exp\left(-\frac{E_2}{kT}\right) + \cdots + \exp\left(-\frac{E_t}{kT}\right)$$

E_j/kT factor: magnitude of E_j relative to kT is the relevant number.

Units of kT: [J] (energy)

Let's look at two limits:

a) T→ ∞ (hi temperature) OR E_j→0 (small energy spacing)

then $\boxed{E_j/kT \rightarrow 0}$

$$p_j^* = \frac{\exp\left(-\frac{E_j}{kT}\right)}{\sum_j^t \exp\left(-\frac{E_j}{kT}\right)} = \frac{1}{(1+1+\ldots 1)} = \frac{1}{t} = p_j^*$$

this means: all states are accessible. Note that

$$Q \rightarrow t$$

b)T → 0 (low temp) OR E_j → ∞ (big energy spacing)

then $\boxed{E_j/kT \rightarrow \infty}$

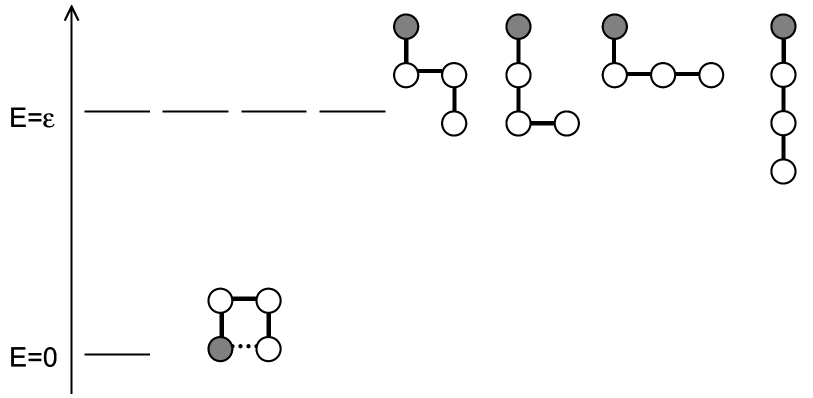
$$p_{j=0}^* = \frac{1}{(1+0+\ldots 0)} = 1 = p_{j=0}^*$$

$$p_{j=\text{rest}}^* = \frac{0}{(1+0+\ldots 0)} = 0 = p_{j=\text{rest}}^*$$

this means: only ground state accessible.

$$Q \rightarrow 1$$

Now let's do it again for our 4 bead polymer:



We still need to account for one more thing:

Degeneracy, g of upper the macrostate--there are four microstates.

$$\boxed{Q \equiv \sum_{l}^{l\,\max} g(E_l)\exp\left(-\frac{E_l}{kT}\right)}$$

l are the levels. $g_{l=0} = 1$, $g_{l=1} = 4$

$$Q = 1\exp(-0) + 4\exp\left(-\frac{\varepsilon}{kT}\right) = 1 + 4\exp\left(-\frac{\varepsilon}{kT}\right)$$

Q(T):

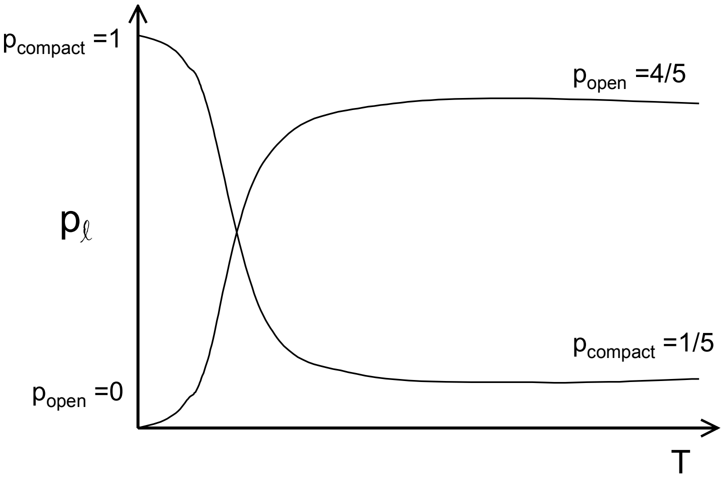
At low T, Q=1 (lowest state accessible)

At high T, Q=5 (all states accessible)

and also p_l (T).

$$p_{\text{compact}} = \frac{1 \cdot e^{\frac{0}{kT}}}{Q} = \frac{1}{Q}$$

$$p_{\text{open}} = \frac{4 \cdot e^{\frac{-\varepsilon}{kT}}}{Q}$$



This is a **unfolding** or a **denaturation profile** for a polymer or protein, etc. Experiments: Fix T, measure p_{open} vs p_{compact}.

- Why are we so interested in Q? We will re-derive thermodynamic properties in terms of Q. This is the **link between the microscopic and macroscopic descriptions**.

Interesting side note: Calculate ΔS of unfolding using S=k ln Ω

$S_{\text{closed}} = k \ln 1 = 0$

$S_{\text{open}} = k \ln 4$

$\Delta S = +$

This says that the protein will want to unfold, based only on entropy. However, this model does not account for things like interaction with the water molecules around the protein, which order around the chain.