

## Lectures on Entropy

After a formal course on thermodynamics most chemistry or biochemistry students remember three formulas.

$$\Delta G^0 = \Delta H^0 - T\Delta S^0$$

$$\Delta G = \Delta G^0 + RT \ln Q$$

and

$$\Delta G^0 = -RT \ln K$$

and that is a very good thing. These three formulas are the center of chemical thermodynamics, and if you do not remember them or where they came from, go back and look them up.

Beyond those formulas most students leave their Pchem courses pretty confused. In particular they are confused about Entropy. Where it comes from and how it influences equilibria. This is not your fault, and it is not Entropy's fault. The problem is simply that most introductory thermodynamics courses use the historical derivation of thermodynamics that relies on macroscopic properties of substances. This version of thermodynamics was developed long before it was clear that there are atoms and molecules made up from those atoms. Thermodynamics was originally derived to explain things like steam engines and to determine the maximal amount of work that could be extracted from those machines. On the other hand, in modern chemistry – in biochemistry in particular- we think about reactions on a molecular basis. When we think about DNA we do not think about a bucket full of one mole of DNA and we are not interested in the amount of heat that is needed to warm this bucket of DNA by one degree C. We are thinking about DNA as a single molecule made up of individual atoms and we want to understand the properties of this single molecule and how it may interact with other molecules. So let's start from scratch and see if we can come up with a picture of thermodynamics that is more useful for our molecular-level view.

The state at which a system consisting of many molecules comes to equilibrium is determined by the battle between two tendencies of every physical system. These tendencies are expressed in the following extremum principles.

- 1) **The principle of minimal potential energy**
- 2) **The principle of maximal multiplicity**

The currency in which the principles of minimal potential energy and maximal multiplicity barter over the exact location of the equilibrium is energy as it is expressed in the equations

$$\Delta G^0 = \Delta H^0 - T\Delta S^0$$

and

$$\Delta G^0 = -RT \ln K$$

At the end of the coming lectures you should have a pretty good idea of how these two extremum principles act at the molecular level and how they are related to enthalpy and entropy and how they influence the equilibrium of every chemical reaction.

### **The principle of minimal potential energy.**

Lets start with the principle of minimal potential energy. This is pretty easy to understand, because this principle of minimal potential energy acts on macroscopic systems just the same way as it does on microscopic systems. Just think of a ball in a mechanical well or two balls connected by a spring. If we leave these systems alone, they will eventually adopt their state of minimal potential energy. This potential energy is directly equivalent to the enthalpy of a system of microscopic particles. The enthalpy of a mole of molecules is simply the sum of the minute potential energies of each of the molecules in the system.

Lets say we have a molecule that likes to form a straight line, if we force this molecule to adopt anything but a straight conformation we need to expend energy and this energy is now stored in these molecules. The tendency for each molecule is now to straighten itself out again (just like a macroscopic spring) and release the energy that we put into the system. In other words the principle of minimal potential energy would predict that all molecules would be perfectly straight.

### **The principle of maximal multiplicity.**

This is the second extremum principle. Unlike the first principle it cannot be explained by a macroscopic equivalent of a single molecule-like object. Instead it is a principle that is statistical in nature. It emerges only when we look at a system of a large number of individual objects. So understanding this principle and how it relates to entropy will take a lot more attention.

Lets think about the simple example of a series of coin tosses. Lets say we toss a coin 10 times. (As a quick reminder if we try to determine the probability of a specific series of independent events, we simply have to multiply the probabilities of the individual events.) So what is the probability to find 10 heads?

HHHHHHHHHH

$$p_h p_h p_h p_h p_h p_h p_h p_h p_h p_h = 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \\ = 0.5^{10} = 0.000976...$$

Wow, that is pretty unlikely, but that sort of makes sense because what are the chances of getting 10 heads in a row?

Now in the back of your head you are thinking that the principle of maximal multiplicity somehow represents entropy. You also vaguely remember that entropy has something to do with probability and that disordered systems are favored by entropy. So what is the probability of finding a disordered series? Lets say:

HTTHHHTHTT

$$p_h p_t p_t p_h p_h p_h p_t p_h p_t p_t = 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \\ = 0.5^{10} = 0.000976...$$

Darn, it is exactly the probability as the all heads case even though it is “disordered” and it even fulfills the statistically expected ratio of 1/2 heads 1/2 tails.

We can keep going with this and we will find that the probability of finding a particular sequence as a result of sequential and independent trials is

$$p(n_1 n_2 n_3 \dots n_i, N) = p_1^{n_1} p_2^{n_2} p_3^{n_3} \dots p_i^{n_i}$$

where

$$\sum_i n_i = N$$

and

$$\sum_i p_i = 1$$

Each specific sequence of heads or tails is just as likely to occur as any other and what we are really interested in is not the probability of a particular series of events but the probability of finding a sequence that has a particular “macroscopic” property. Macroscopic, in this context, means that we do not care what the order of the heads and tails is, but simply how many of the coins are heads or tails. This sort of question involves multiplicity rather than probability.

## Multiplicity

So what is multiplicity? The Multiplicity of a set of combined events is the number of different ways in which this set of events could possibly occur. For a simple example, lets say a car company makes two models of cars and paints them in three different colors. The multiplicity of the “event” called car then is 6.

To better understand multiplicity lets think of a sequence of coin tosses. So lets say we flip a coin four times or we have four coins lying on a table and we pick each one up and flip it once. When we do this we will often get an outcome where half of the flips gave us a head and half gave us a tail, but sometimes will also get all heads or all tails. To see, if we can understand the relatively low likelihood of seeing all heads compared to seeing 2 heads and 2 tails in terms of multiplicity, lets write down all the different ways in which we can get all heads, 3 heads 1 tail etc. etc.

0 T: HHHH	1 way
1 T: THHH, HTHH, HHTH, HHHT	4 ways
2 T: TTHH, HTTH, HHTT, THHT, HTHT, THTH	6 ways
3 T: HTTT, THTT, TTHT, TTTH	4 ways
4 T: TTTT	1 way

We can see two things. First the multiplicity of the two perfectly ordered results (all heads and all tails) has a lower multiplicity than sequences with mixed results. In other words there are a lot more microscopic ways to have 2 heads and 2 tails up than there are ways to have all heads or all tails. So the chances that we will end up with a system that shows 2 heads and 2 tails is higher than the chance of seeing all heads. We also see that the case of 50% heads and 50% tails turns out to be the one with the highest multiplicity, reconfirming our intuition that the “statistically expected” outcome is the highest multiplicity case.

Writing things out in this way works pretty well for a small number of  $N$ , but of course it would get a little cumbersome, if we had to write this out for  $N_A$  molecules for a chemical reaction. So let's look for a general equation to determine the multiplicity of a specific macroscopically distinguishable state.

Let's say you have a jar of coins and each coin is numbered somewhere on its side. Let's take out one coin at a time and arrange them in a sequence. It's kind of silly to point this out, but if you only have one coin there is only one way in which you can arrange this coin. If you have two coins then there are  $1 \cdot 2$  ways to arrange them, one coin first and the other second or the second coin first and the first coin second. Adding a third coin, you can generate three new arrangements for each of the two arrangements made in the two coin case by placing the third coin either before coin 1, between coin 1 and 2 or after coin 2 resulting in a multiplicity of  $1 \cdot 2 \cdot 3$ . We can keep going following the same principle and we will find that multiplicity  $W$  of number  $N$  of outcomes can be calculated with the following simple formula.

$$W(N) = N!$$

*where*

$$N! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot \dots \cdot N$$

*remember*

$$0! = 1$$

This is the multiplicity of a system if we can actually distinguish between each of the objects. But in the case of a test tube full of molecules, we cannot distinguish each molecule, instead we can only observe if a molecule is in state A or state B – heads or tails.

To better see this  
Consider just the case of three heads and one tail

If we could distinguish between each head we would have the following increase in possibilities for the first case

And we would have the same number of different arrangements for each one of the other instances of having one tail up. By dividing the multiplicity we would get if all coins were distinguishable by the multiplicities we would get if all the heads were distinguishable and all the tails were distinguishable we get the multiplicity of being able to distinguish only between a coin being heads or tails.

In General the multiplicity of a system with  $N$  different objects where  $n_1$  objects have property 1 and  $n_2$  objects have property 2 etc. is given by.

$$\sum_{i=1}^t n_i = N$$

## Understanding the formula for multiplicity

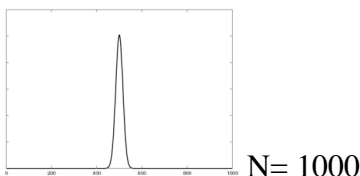
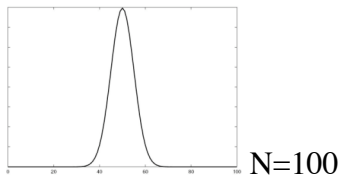
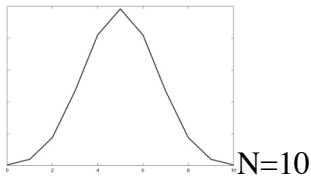
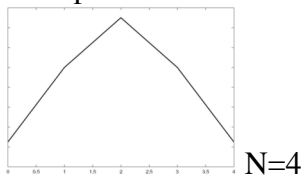
To understand the equation for multiplicity a bit better let's play with it a bit.

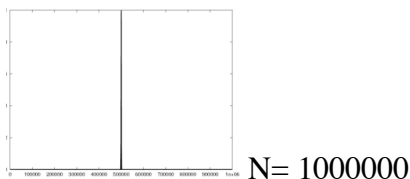
-Example 1: A very simple case: If we have  $n_1=N$ , in other words we always get the same result (e.g. all heads). Then  $W$  computes to  $N!/N!=1$ . There is only one microscopic way in which we can get this macroscopic result. That is exactly the result we would expect intuitively.

-Example 2: One tail between all heads If we have only a single result that is different,  $n_1=N-1$  and  $n_2=1$ , any one of the  $N$  coins could be the one that is “tails” so we have  $N$  different microscopic ways to achieve the same macroscopic result. Plugging in the numbers, we get exactly this result back  $N!/(N-1)! * 1! = N$ .

-Example 3:  $W$  is maximal if we have equal numbers of heads and tails. Just by looking at the formula it should be clear that  $W$  will be maximal if  $n_1=n_2$ . For example for a system with 4 members  $4!/2! \times 2! = 6$  while  $4!/3! \times 1!$  is 4.

-Example 4: The  $W$  distribution as a function of  $N$  Here are normalized plots of  $W(n, N-n; N)$  as a function of  $n$  for various  $N$ . In other words, each plot shows the relative multiplicities for different ratios of heads and tails.





As you see the distribution of populations is always centered on the point where you get 1/2 heads and 1/2 tails. What you also notice is that the function becomes very narrow for large  $N$ . This means that the multiplicity of a systems with large  $N$  (as most molecular systems) drops precipitously as we move just a little bit away from the most probable distribution of 50% heads 50% tails. In other words it becomes very unlikely to ever find any case in which we flip 1000000 coins (let alone  $N_A$  molecules) and get any distribution that even remotely deviates from the statistically expected result. So while the flipping process is perfectly random at the level of the individual event, the predictions we make for a large combination of events is practically deterministic. This tendency of a large population to adopt states that perfectly match the statistically predicted ratios of outcomes is called the **principle of fair apportionment**.

-Example 4:  $W$  as a function of  $t$  Lets now compare two series of six events each, both of which have an even distribution between their outcomes but lets vary  $t$ , the number of distinguishable outcomes for each event. In other words lets compare our coins with objects that have three sides (there are no simple objects with three sides by the way, but lets pretend they did exist).

If we distribute our trial results over three different outcomes  $W = 6! / 2! * 2! * 2! = 90$ , but if we have only two distinguishable outcomes  $W = 6! / 3! * 3! = 20$ . We see that the multiplicity of a system also increases as the number of distinguishable results increases.

Our challenge now is to convert this principle that  $W$  should be maximized into an energy that we can then use in the calculation of  $\Delta G$ .

One thing should strike you as funny in this respect. Consider the example above, but now with twice as many molecules. If  $W$  is the multiplicity of a system with  $N$  objects, the multiplicity of a system with  $2N$  objects is  $W^2$ . In other words the multiplicity per object is a function of the number of objects. This is clearly different from the way entropy works. If we get a  $\Delta S$  of 5cal/K for converting one mol of A into B then doing this for 2 mol will give 10cal/K. So something strange is going on in the relationship of entropy and multiplicity.

### **From Multiplicity to Entropy**

Yesterday we thought about a number of systems that were pretty cartoonish which demonstrated how the principle of maximal multiplicity explains at least qualitatively some macroscopic events that we observe in everyday life. The question now is how we can convert this multiplicity into an energy in kcal/mol. It turns out that establishing the connection was a seriously difficult problem and I will not attempt to re-derive it here. Fortunately this problem was solved for us many years ago by Ludwig Boltzman and even better, it takes a remarkably simple form.

To get to the Boltzmann equation lets do a little example calculation that is closer to a real problem. Lets say we have a mole of tetrahedrons that has three black sides and one white side. To begin with they all face with the white side down, now how much Multiplicity would we gain if we were to run a reaction so that at the end all the tetrahedron's have a black side facing down.

$$W(n_{white} = N_A, n_{black1} = 0, n_{black2} = 0, n_{black3} = 0; N_A) = \frac{N_A!}{N_A! 0!} = 1$$

$$W(n_{white} = 0, n_{black1} = \frac{N_A}{3}, n_{black2} = \frac{N_A}{3}, n_{black3} = \frac{N_A}{3}; N_A) = \frac{N_A!}{\frac{N_A!}{3} \frac{N_A!}{3} \frac{N_A!}{3}}$$

Now I admit that we are a little bit sloppy here because we are considering only the possibility that the three black sides are present with equal probability, but as we have seen before, with large numbers, the fair apportionment principle is a fairly strong “force” and so we are justified in writing the equation down this way.

*(As it turns out, this neglect of other configurations such as*

$$\frac{N_A!}{\left(\frac{N_A}{3} - 2\right)! \left(\frac{N_A}{3} + 1\right)! \left(\frac{N_A}{3} + 1\right)!}$$

*or*

$$\frac{N_A!}{\left(\frac{N_A}{3} + 3\right)! \left(\frac{N_A}{3} + 1\right)! \left(\frac{N_A}{3} - 4\right)!}$$

*is perfectly compensated for by Stirling's approximation we make below. Stirling's approximation generally underestimates  $N!$  and this mistake is larger for smaller  $N$ . What is happening is that the nominator with its smaller numbers is under estimated by the Stirling's approx to a larger extent than the bigger numerator. This is the same effect as neglecting the broader distribution of  $N$  among additional microscopic conformations, which also leads to a smaller numerator. In this sense Stirling's approximation is a truly sterling approximation. Sorry, I could not resist the pun!)*

Factorials of numbers of the order of  $N_A$  are very difficult to handle (my calculator will give out after about  $60!$ ), but there is a mathematical trick that we can use. It is called the Sterling approximation and it says that:

for very large  $x$

$$x! = \left(\frac{x}{e}\right)^x$$



so

$$W(n_{white}=0, n_{black1}=\frac{N_A}{3}, n_{black2}=\frac{N_A}{3}, n_{black3}=\frac{N_A}{3}; N_A) = \frac{\left(\frac{N_A}{e}\right)^{N_A}}{\left(\frac{N_A}{3e}\right)^{\frac{N_A}{3}} \left(\frac{N_A}{3e}\right)^{\frac{N_A}{3}} \left(\frac{N_A}{3e}\right)^{\frac{N_A}{3}}}$$

$$= \frac{\left(\frac{N_A}{e}\right)^{N_A}}{\left(\frac{1}{3}\right)^{N_A} \left(\frac{N_A}{e}\right)^{N_A}} = 3^{N_A}$$

We have  $N_A = 6 \cdot 10^{23}$  in the exponent, this is obviously a huge number. So there are a tremendous number of microscopic ways to get the macroscopic result of having only black surfaces facing down and there is only one way of having all surfaces white. But how much is this big discrepancy between the multiplicity of the all-white-down and the all-black-down state of the system worth in kcals/mol.

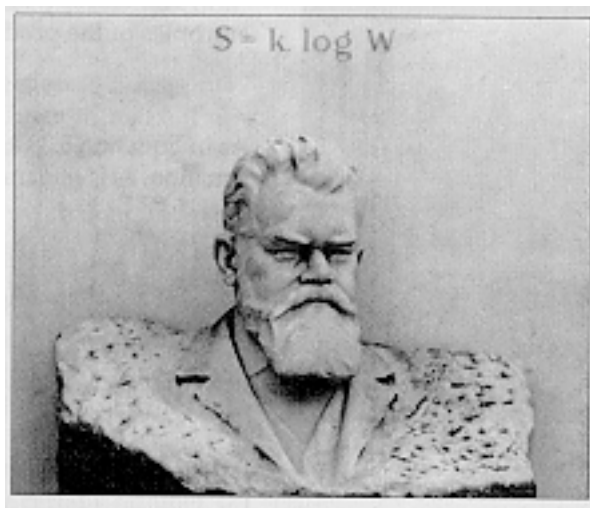
### The Boltzmann equation

With his famous Equation Ludwig Boltzmann established the link between a property of the microscopic world (multiplicity) and the macroscopic world (energies, entropy and equilibrium constants), which we can observe in a laboratory. His equation allows us to take any molecular model and -from it- predict a broad range of experimentally observable macroscopic results. And these calculations start from scratch, no empirical parameters no fitting, etc.

This feat is even more remarkable, since during Boltzmann's time the existence of atoms and that those atoms came together to form molecules was not generally accepted.

So to make sure that nobody ever lost his precious equation, he put it on the tombstone of his grave in the Zentralfriedhof in Vienna, Austria.

I am not kidding



In those days they used log to specify the log<sub>e</sub>, which we today write as “ln” So the modern day version of the Boltzmann Equation is:

$$S = k \ln W$$

Where W is the multiplicity and k is the Boltzmann constant ( $0.33 \times 10^{-23}$  cal/K) which emerged from his calculations.

**Just looking at this equation should answer our earlier question why the effect of doubling the size of a reaction only leads to a doubling of that reaction's entropy. Doubling the number of molecules in a reaction results in a multiplicity that is the square of multiplicity of the two separate reactions. However, since the entropy is proportional to the ln of the multiplicity and doubling the number of the molecules leads only to a doubling of the entropy.**

$$\ln(a \cdot b) = \ln a + \ln b$$

Lets apply the Boltzmann equation to our tetrahedron case and see if we can calculate a standard free energy for our reaction.  
Since the enthalpic contribution to our reaction all-white bottom -> all-black bottom is zero.

$$\begin{aligned} \Delta G^0_{white \rightarrow black} &= -TS^0_{black} - (-TS^0_{white}) \\ &= -T \cdot k \cdot \ln W_{black} - (-T \cdot k \cdot \ln W_{white}) \\ &= -T \cdot k \cdot \ln W_{black} - 0 \\ &= -298 \text{ K} \cdot 0.33 \cdot 10^{-23} \text{ cal/K} \cdot \ln 3^{N_A} \\ &= -298 \text{ K} \cdot 0.33 \cdot 10^{-23} \text{ cal/K} \cdot N_A \cdot \ln 3 \\ &= -298 \text{ K} \cdot 0.33 \cdot 10^{-23} \text{ cal/K} \cdot 6.022 \cdot 10^{23} \text{ mol}^{-1} \cdot 1.099 \\ &= -650 \text{ cal/mol} = -0.65 \text{ kcal/mol} \end{aligned}$$

Since we know that entropy is an *extensive* property, we could have actually calculated the entropy of a single molecular tetrahedron and then added up this molecular entropy over Avogadro's number of tetrahedrons.

As it turns out this calculation is extremely easy.

The tetrahedron has four faces and we will assume that the first one is the one that is facing down. There are now four ways to arrange the faces.

B B B W	downward face black
B B W B	downward face black
B W B B	downward face black
W B B B	downward face white

In other words the multiplicity of being black is 3 and the multiplicity of being white is 1. With Boltzmann's equation we then get:

$$\begin{aligned}
 S_{black} &= k \ln 3 \\
 &= 0.33 \cdot 10^{-23} \text{ cal/K} \cdot 1.099 \\
 &= 3.6 \cdot 10^{-24} \text{ cal/K}
 \end{aligned}$$

and

$$S_{white} = k \ln 1 = 0$$

Then multiply the molecular entropy by Avogadro's number.

$$\begin{aligned}
 S_{black}^0 &= N_a (k \ln 3) \\
 &= 6.022 \cdot 10^{23} \text{ mol}^{-1} \cdot 0.33 \cdot 10^{-23} \text{ cal/K} \cdot 1.099 \\
 &= 2.184 \text{ cal/K} \cdot \text{mol}
 \end{aligned}$$

and

$$\begin{aligned}
 \Delta G_{white \rightarrow black}^0 &= -TS_{black}^0 \\
 &= -298 \text{ K} \cdot 2.18 \text{ cal/K} \cdot \text{mol} = -650 \text{ cal/mol} = -0.65 \text{ kcal/mol}
 \end{aligned}$$

We get the same answer if we use the good old formulas:

$$\Delta G^0 = \Delta H^0 - T\Delta S^0$$

and

$$\Delta G^0 = -RT \ln K_{equil.}$$

since our model says that there is no enthalpic effect:

$$\Delta G^0_{white \rightarrow black} = -T\Delta S^0_{white \rightarrow black}$$

so

$$-T\Delta S^0_{white \rightarrow black}$$

$$= -RT \ln K_{equil.}$$

$$= -RT \ln \frac{[black]}{[white]}$$

$$= -RT \ln \frac{3}{1}$$

$$\text{with } R = k \cdot N_A$$

$$= -T \cdot k \cdot N_A \cdot \ln 3$$

See that we get exactly the same result as we did with the calculation based on our molecular model.