

# Lectures on Entropy and free Energy

## Lecture 1

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

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

and

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

and that is a very good thing. These two 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 two 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.

Fortunately there is an approach to thermodynamics that takes this same molecular view of the world. This type of thermodynamics is called statistical thermodynamics and Ludwig Boltzmann is its father and we will later see and use his famous formula for entropy.

So lets start from scratch and see if we can arrive a view of thermodynamics that is more appropriate for our type of problems. The state at which a system consisting of many molecules comes to an equilibrium is determined by the battle between two tendencies of every physical system. These tendencies are expressed in the following two extremum principles.

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



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

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

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 event 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.

In our case we want to know what is the multiplicity of sequences of coin tosses that results in a given number of tails. Lets see if we can find and make sense of a general formula to calculate multiplicity. Lets think about the coin problem again.

Lets say you have a jar of coins and each coin is numbered somewhere on its side. Lets take out one coin at a time and arrange them in a sequence:

Its 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

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

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 the properties of the whole population of molecules: What fraction of coins shows tails, so to speak?

So lets say we flip a coin four times. How many different microscopic ways are there to obtain the same macroscopic result. As an example let us calculate this for random tosses of four coins:

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. We also see that the case of 50% heads and 50% tails turns out to be the one with the highest multiplicity.

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 in a 1 molar reaction. So lets look for a general equation to determine the multiplicity of a specific macroscopically distinguishable state.

The formula we try to find has to correct the multiplicity of a hypothetical system in which each component can be distinguished for the fact that we can distinguish only between those components that have macroscopically distinguishable properties.

This time let me just give you the general formula and then lets see if we can understand it. For a case of a system with N components which can adopt t different properties and where  $n_i$  is the number of times a particle adopts property i we get the following formula for the multiplicity W :

$$W(n_1, n_2, n_3, n_4, \dots, n_t, N) = \frac{N!}{n_1! n_2! n_3! n_4! \dots n_t!}$$

where

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

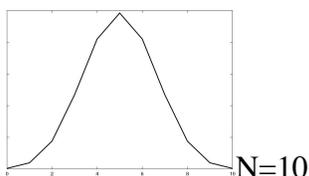
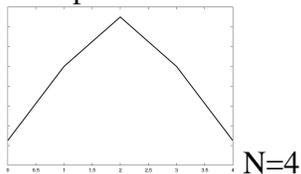
### Understanding the formula for multiplicity

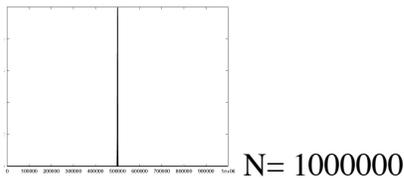
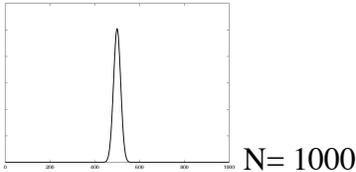
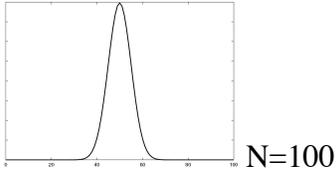
To understand the equation for multiplicity a bit better lets play with it a bit.

-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: 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.

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.

## **Lecture 2**

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.

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  $x$

$$x! = \frac{x^x}{e^x}$$

so

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

$$= \frac{\frac{N_A^x}{e^x}}{\frac{1}{3} \frac{N_A^{\frac{N_A}{3}}}{e} \frac{N_A^{\frac{N_A}{3}}}{e} \frac{N_A^{\frac{N_A}{3}}}{e}} = 3^{N_A}$$

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 and the all black down state of the system worth in kcals/mol.

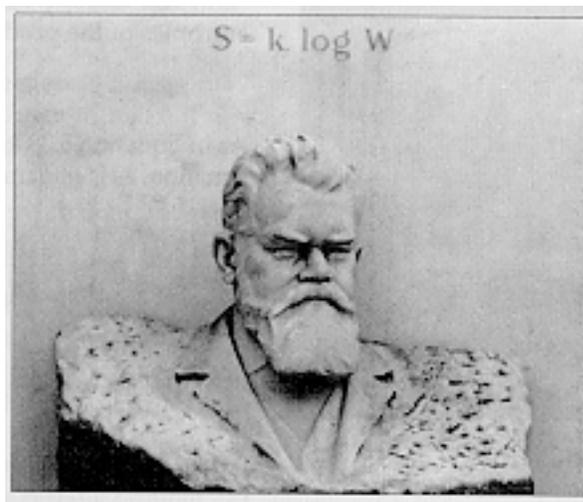
### The Boltzmann equation

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  $\log_e$ , 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.

Why ln and not any other parameter, -> This makes entropy extensive  
 $\ln(a \cdot b) = \ln a + \ln b$

at Room temperature if there is no enthalpic contribution to this reaction the molar free energy for the reaction of an all-white bottom  $\rightarrow$  all-black bottom is.

$$\Delta G^0_{white \rightarrow black} = \Delta TS^0_{black} - (\Delta TS^0_{white}) = \Delta 300 \text{ K} \cdot 2.18 \text{ cal/K} \cdot \text{mol} = 0.64 \text{ kcal/mol}$$

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:

$$S_{black} = k \ln 3 = k 1.099 = 3.6 \cdot 10^{24} \text{ cal/K}$$

and

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

Since entropy is an *extensive* property we simply multiply the molecular entropy by Avogadro's number.

$$S^0_{black} = N_a (k \ln 3) = 6 \cdot 10^{23} \text{ mol}^{-1} \cdot 3.6 \cdot 10^{24} \text{ cal/K} = 2.18 \text{ cal/K} \cdot \text{mol}$$

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:

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

so

$$\Delta T \Delta S^0_{white \rightarrow black} = RT \ln K_{equil.} = RT \ln \frac{[black]}{[white]} = RT \ln \frac{3}{1} = TN_A (k \ln W_{black} - k \ln W_{white})$$

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

Here finally is an expression of Entropy that is even easier to handle when dealing with molecular models.

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

$$W = \frac{N!}{n_1! n_2! \dots n_t!}$$

sterling

$$W = \frac{(N/e)^N}{(n_1/e)^{n_1} (n_2/e)^{n_2} \dots (n_t/e)^{n_t}}$$

$$= \frac{N^N}{n_1^{n_1} n_2^{n_2} \dots n_t^{n_t}} = \frac{1}{p_1^{n_1} p_2^{n_2} \dots p_t^{n_t}}$$

take ln

$$\ln W = - \sum_{i=1}^t n_i \ln p_i$$

$$\frac{\ln W}{N} = - \sum_{i=1}^t p_i \ln p_i = \frac{S_N}{Nk}$$

and

$$\frac{S_N}{Nk} = \frac{S}{k}$$

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

## Summary

Entropy is not mysterious. Entropy simply reflects the fact that there are many more microscopic ways to achieve certain macroscopic results. Hopefully we also have replaced that foggy relationship between entropy and probability with a clear and easy to understand quantitative relationship. Entropy is maximal if we have a flat distribution of probabilities, and if there are no other constraints on a system, a distribution will always

adopt the flattest possible probability distribution. If a system adopts anything other than the flat distribution, then this is a sign of bias in our system.

Even though entropy is statistical in Nature, the free energy derived from purely entropic processes delivers real energy, the kind of energy that allows you to run a light bulb or –since we are in a biochem course- that makes ATP.