

Chemistry 21b – Spectroscopy & Statistical Thermodynamics

Lecture # 22 – Ensembles & the Boltzmann Law

As stated in the supplementary Thermodynamics Primer to Lecture 21, the central goal of statistical thermodynamics is to calculate the thermodynamic properties of a substance in terms of the molecular properties we have learned about in Ch21a and b. That is, given the structure of the individual molecular components of a sample and an intermolecular potential that governs their interactions with each other, we would like to be able to calculate from first principles thermodynamic quantities such as entropy, free energy, and the like. How might we do this?

From a macroscopic viewpoint, we can characterize a mole of liquid water with only a few parameters – volume, concentration or density, temperature, etc. As we'll show below, however, the quantum mechanical degeneracy of an isolated N -body system is of order 10^N . Characterizing each and every microscopic state of such a system is clearly impossible! Fortunately, as it turns out what is required is the calculation of the average property under the assumption that each possible quantum state has the same weight, or probability; where we *postulate* that the average calculated property from a microscopic point of view corresponds to a parallel thermodynamic property (average energy = thermodynamic energy, average pressure = thermodynamic pressure, etc.). In order to calculate these averages, it will first be useful to remind ourselves of a few things about probability distributions, especially as the number of particles approaches infinity.

Stirling's Approximation

In statistical thermodynamics, the factorials of very large numbers (Avogadro's number to be more precise) are often involved. This can become quite tedious, and so it is very helpful to have an approximation of $N!$ for large N . These are called asymptotic approximations in mathematics. Here, it will be more useful to deal with $\ln(N!)$ because this is a sum rather than a product.

Since $N! = N(N-1)(N-2)\cdots(2)(1)$, $\ln(N!) = \sum_{m=1}^N \ln(m)$. A plot of $\ln(x)$ versus x , where x is an integer, is shown by the histogram in Figure 22.1. The continuous curve of $\ln(x)$ is also plotted. For the integer approach, the sum of the areas in the histogram up to N is $\ln(N!)$. Clearly, the continuous form of $\ln(x)$ forms an envelope to the histogram, and the envelope becomes a steadily smoother approximation as x increases. The largest errors in the calculation of the area are seen to occur at small values of x . If N becomes large enough (remember we're thinking about deriving an asymptotic expansion), the error at small values will make a negligible contribution to the *total* area. Thus,

$$\ln(N!) = \sum_{m=1}^N \ln(m) \approx \int_1^N \ln(x) dx = N \ln(N) - N \quad (N \text{ large}) \quad (22.1)$$

which is Stirling's approximation to large N . The lower limit could actually be taken as 0 since $x \ln(x) \rightarrow 0$ as $x \rightarrow 0$. A somewhat more accurate approximation is

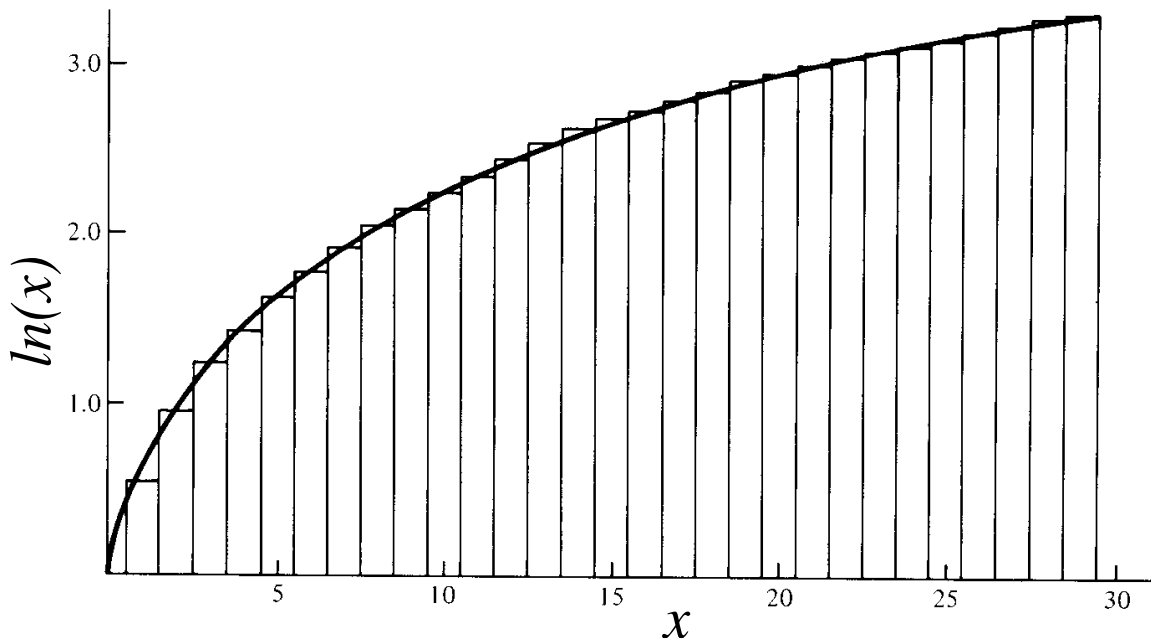


Figure 22.1– A plot of $\ln(x)$ versus x , showing how the summation of $\ln(m)$ can be approximated by $\int \ln(x)dx$.

$\ln N! \approx N \ln(N) - N + \ln(2\pi N)^{1/2}$, but this last term can be safely ignored when dealing with a macroscopic sample of atoms or molecules.

Multinomial Distributions for Large N

We will have several occasions in what follows to consider the problem of determining how many ways it is possible to divide N *distinguishable* systems into groups such that there are N_1 systems in the first group, N_2 systems in the second group, and so on such that $N_1 + N_2 + \dots = N$. Fortunately, this is one of the simpler problems in combinatorial analysis. First we need to calculate how many permutations there are of the N distinguishable objects, that is, how many ways can we arrange, or order, N distinguishable objects? If we place one of the objects into position one, there are $N - 1$ objects left for position two, and so on until only one object is left for position N . Clearly, the total number of ways of doing this is just the product of all the choices, or $N(N - 1)(N - 2) \dots (2)(1) \equiv N!$ for distinguishable objects. We first consider breaking up the total suite into two groups.

The first group has N_1 objects, the second $N_2 = (N - N_1)$. There are $N(N - 1) \dots (N - N_1 + 1)$ ways to form the first group, and $N_2! = (N - N_1)!$ ways to form the second group, which you might expect would yield a total number of

$$N(N - 1) \dots (N - N_1 + 1) \times (N - N_1)! = \frac{N!}{(N - N_1)!} \times (N - N_1)! = N!$$

from the product. However, the order in which we place N_1 members in the first group and N_2 members in the second group is immaterial to the problem, and so the above equation would over count the combinations. In fact, all $N_1!$ orders of the first group and $N_2!$ orders

of the second group correspond to just one division of N objects into N_1 and N_2 objects. The desired result is therefore

$$\frac{N!}{N_1!(N - N_1)!} = \frac{N!}{N_1!N_2!} \quad , \quad (22.2)$$

and is called the binomial coefficient since it occurs in the expansion of $(x + y)^N$. We have already encountered this formula in our calculation of the fine structure spin-spin splittings in proton NMR in Lecture 19. The generalization of eq. (22.2) to the division of N objects into r groups is straightforward, and is given by

$$\frac{N!}{N_1!N_2! \cdots N_r!} = \frac{N!}{\prod_{j=1}^r N_j!} \quad . \quad (22.3)$$

These multinomial coefficients occur in the expansion of $(x_1 + x_2 + \cdots + x_r)^N$.

What happens when N becomes very large? Let's consider the binomial case for simplicity. In order to derive the most likely outcome, that is the value of N_1 for which the probability distribution $f(N_1) = N!/N_1!(N - N_1)!$ reaches its maximum value, we will treat N_1 and N as continuous variables. Further, since $\ln(x)$ is a monotonic function of x , we'll maximize $\ln[f(N_1)]$. By using Stirling's approximation and setting $[(d \ln f(N_1))/dN_1] = 0$, the value of N_1 which maximizes $f(N_1)$ is found to be $N_1^* = N/2$. What happens if we expand $\ln[f(N_1)]$ about this point? The Taylor expansion is

$$\ln[f(N_1)] = \ln[f(N_1^*)] + \frac{1}{2} \left(\frac{d^2 \ln[f(N_1)]}{dN_1^2} \right)_{N=N_1^*} (N_1 - N_1^*)^2 + \dots \quad ,$$

where the linear term is missing because the first derivative of $\ln[f(N_1)]$ is zero at $N = N_1^*$. The second derivative in the expansion turns out to be $-4/N$. Thus, the expansion can be cast in terms of a Gaussian distribution function, or

$$f(N_1) = f(N_1^*) \exp \left\{ -\frac{2(N_1 - N_1^*)^2}{N} \right\} \quad ,$$

where the mean is $N_1^* = N/2$ and the standard deviation σ is of order $\sigma \sim N^{1/2}$.

For $N \sim 2 \times 10^{20}$, the binomial distribution is a bell shaped curve whose width is only $\sigma \sim 10^{10}$. Put another way, if we normalize this distribution to lie between $0 \leq x \leq 1$, that is, let $x = N_1/N$, the peak of the distribution stays at 0.5, but its relative width becomes smaller and smaller since $\sigma_{norm} = N^{1/2}/N = N^{-1/2}$.

The peak therefore becomes sharper as N (and the N_j) increases, and approaches a delta function in the limit that $N_j \rightarrow \infty$ for all j . Thus, while the number of possible states of a system becomes impossibly large to keep track of for numbers that are relevant to the thermodynamics of macroscopic systems, the probability of large *excursions* from the most likely state becomes increasingly – indeed vanishingly – small. It is this fact that will permit us to calculate thermodynamic properties from a collections, or ensembles, of \sim Avogadro's number of quantum mechanical systems, a topic we shall turn to next.

The Most Probable Distribution for Quantum Systems

Several times in this course we have used the Boltzmann formula to calculate the relative populations of two quantum states i and j of a molecule, or $\frac{N_i}{N_j} = \frac{g_i}{g_j} e^{-(\epsilon_i - \epsilon_j)/kT}$, where the degeneracy and energy of state i are denoted g_i and ϵ_i .

The statistical thermodynamics route to this fundamental relationship begins with a consideration of an ensemble of systems. An ensemble is a collection of a very large number of systems, each constructed to be a replica on a macroscopic level of the particular thermodynamic system of interest. An isolated system with N molecules, a volume V , and energy E fixed is called the *microcanonical ensemble*. The individual energy levels of the particles are specified as $\epsilon_1, \epsilon_2, \dots$, which have populations of N_1, N_2, \dots . Suppose the entire ensemble is made up of A systems, and is surrounded by thermal insulation, thus making it an isolated system with volume $\mathcal{V} = AV$, number of molecules $\mathcal{N} = AN$, and total energy \mathcal{E} . This *canonical ensemble* is one for which \mathcal{N} , \mathcal{V} , and T (or the average energy of a system in the ensemble) are specified. As $N \rightarrow \infty$, there are obviously an extremely large number of combinations of molecules in distinct eigenstates (these specific combinations are called microstates) that are compatible with the microcanonical energy E . For now, we will assume that each particle is *distinguishable* from all the others and that there are no restrictions on how the particles may be assigned to the various energy levels. In order to calculate the average energy and pressure for the ensemble, we will need the two postulates of statistical mechanics:

First Postulate: The time average of a mechanical property for a single system is equal to the average of this property over all the systems of the ensemble.

Second Postulate: In an isolated system all of the possible quantum states consistent with specified values of the number of particles, the volume, and the energy are equally probable. This is the *principle of equal a priori probabilities*.

The probability of any particular distribution is by definition equal to the number of ways of realizing that distribution, characterized by a particular set of occupation numbers N_j , is by definition equal to the number of ways of realizing that distribution. If we define this distribution as $\Omega(E)$, from the discussion above the form of this distribution is

$$\Omega(E) = \frac{N!}{\prod_j N_j!} \quad (22.4)$$

The equilibrium distribution is the one for which $\Omega(E)$ is a maximum. Fortunately, there are two constraints we can apply to further define the problem – the constancy of the number of particles and the constancy of the total energy, or

$$\sum_j N_j = N \quad \text{and} \quad \sum_j N_j \epsilon_j = E \quad (22.5)$$

To maximize $\Omega(E)$, we must find the conditions under which its variation is zero for infinitesimal changes, or $d\Omega = 0$. By taking the \ln of both sides of eq. (22.4) and recognizing that the constraints imply that

$$dN = \sum_j dN_j = 0 \quad \text{and} \quad dE = \sum_j \epsilon_j dN_j = 0 \quad (22.6)$$

(since N and E are constant for the ensemble), we find

$$d \ln \Omega = \sum d \ln N_j! = \sum \ln N_j dN_j = 0 \quad (22.7)$$

using Stirling's approximation.

At this point in the analysis it is common to use the *method of Lagrange multipliers*, which states (see Apostol's *Calculus, Vol. II*): If a scalar field $f(x_1, \dots, x_n)$ has a relative extremum when it is subject to m constraints, say

$$g_1(x_1, \dots, x_n) = 0, \quad \dots, \quad g_m(x_1, \dots, x_n) = 0,$$

where $m < n$, then there exist m scalars $\lambda_1, \dots, \lambda_m$ such that

$$\nabla f = \lambda_1 \nabla g_1 + \dots + \lambda_m \nabla g_m$$

at each extremum point. In practice the system of $n + m$ equations obtained by taking the m constraints and n scalar equations are solved for the $n + m$ unknowns x_1, \dots, x_n and $\lambda_1, \dots, \lambda_m$. The points (x_1, \dots, x_n) at which relative extrema occur are found among the solutions to these equations.

The scalars $\lambda_1, \dots, \lambda_m$ that are introduced to help solve the problem are called *Lagrange's multipliers*, and one multiplier is introduced for each constraint (we have seen a version of this approach in the first week's lectures on perturbation theory). Here, there are two constraints, and hence two multipliers that we will call α and β . Applying these to eq. (22.6) and combining them with eq. (22.7) yields

$$\sum_j \alpha dN_j + \sum_j \beta \epsilon_j dN_j + \sum_j \ln N_j dN_j = 0. \quad (22.8)$$

The variations dN_j may now be considered to be perfectly arbitrary (the constraining conditions having been removed). Thus, for eq. (22.8) to hold, *each term in the summation must vanish*; and so

$$\ln N_j + \alpha + \beta \epsilon_j = 0$$

or

$$N_j = e^{-\alpha} e^{-\beta \epsilon_j}. \quad (22.9)$$

As we have seen in our analysis of molecular spectra, it is common for there to be situations in which there is more than one state that corresponds to a given energy level ϵ_j ; and this is called the degeneracy g_j . Thus, the more general form of eq. (22.9) is

$$N_j = g_j e^{-\alpha} e^{-\beta \epsilon_j}. \quad (22.10)$$

Since $N = \sum_j N_j$, we can write

$$N = \sum_j N_j = e^{-\alpha} \sum_j g_j e^{-\beta \epsilon_j} = e^{-\alpha} q,$$

where $e^{-\beta\epsilon_j}$ is the Boltzmann factor and

$$q = \sum_{j(\text{levels})} g_j e^{-\beta\epsilon_j}$$

is called the microcanonical partition function. It should look familiar to you from our discussion of rotational and vibrational spectra. The expression for the occupation number of a given level N_j therefore becomes

$$\frac{N_j}{g_j} = \frac{N}{q} e^{-\beta\epsilon_j} . \quad (22.11)$$