ENSEMBLES AND ALL THAT

Physics 664 Statistical Mechanics

Ensembles are mental constructs that contain replicated systems chosen in such a way as to mimic real systems. There are different kinds of ensembles depending on the nature of the systems of interest, but all of which have much in common. We will begin by developing the formalism for the ***Canonical Ensemble*** because it illustrates the mechanics of the problem at hand and is mathematically the most tractable. However, it may be helpful to first contrast the more elementary but less abstract method often used in introducing statistical thermodynamics for the first time in the typical undergraduate course on the subject. The following table compares the 'System/Ensemble' method of J. Willard Gibbs (1902) with the more elementary 'Particle/System' method of Boltzmann, 1877.

COMPARISON OF THE ENSEMBLE/SYSTEM FORMALISM VS THE SYSTEM/PARTICLE FORMALISM

ENSEMBLE/SYSTEM FORMALISM	**SYSTEM/PARTICLE FORMALISM**
Stirlings formula is exact as $N_{SYSTEMS} \rightarrow \infty$	$N_{PARTICLES} << \infty$ so Stirling's formula is approximate
Systems are macroscopic so labels are okay	Real particles do not come with labels
Systems can be composed of strongly interacting particles	This formalism assumes weak interactions between particles
S = k log Ω(N,V,E) follows logically from the mathematics and thermodynamics	S = k log Ω(N,V,E) must be postulated (Boltzmann, 1877)
Harder to describe the 'single particle' partition function since the system partition is the natural identity	The single particle partition function is the natural identity

First, some definitions:

System:

A hypothetical enclosure of volume V containing N particles at temperature T that is constructed (mentally) to replicate a real system of interest. The walls of the system are diathermal and closed. The system has an associated energy state spectrum $\{E_i\}$ that is some combination of the constantly fluctuating states of the individual particles making up the system. We make no assumptions at this point regarding the nature of the particles, their state or the nature of how they interact – strongly, weakly or not all. Because the particle energy states making up the system change constantly due to collisions, interactions and the natural time evolution of the particle wave functions, the specific energy state of the system will also change with time. What does not change is the *spectrum* of possible energies available to the system (the system is isothermal), although again, we cannot know those details either. Consequently, fluctuations can occur over time in the mechanical variables like the energy and pressure, but V, N and T remain fixed in our system.

Canonical Ensemble:

Now consider an adiabatically sealed structure (hence isolated) consisting of N_T ($N_T \rightarrow \infty$) identical systems of the type described above, and all of which are embedded in a common heat bath and are therefore in thermal equilibrium, i.e., the systems in the ensemble are isothermal. All systems are defined by exactly the same T, V, N and energy state spectrum, $\{E_i\}$. Furthermore, we assume that the thermal interactions between systems do not influence their energy state spectrums. Thus, T, V, N and $\{E_i\}$ are the same for all N_T systems making up the ensemble. Finally we assume the systems are macroscopic and can be labeled.

Constraints:
Since the ensemble is isolated, it will be characterized by a fixed energy $E_T = \sum n_i E_i$ where E_T = constant. We can also speak of the number of systems, n_i, in the i^{th} **system** energy state where $\sum n_i = N_T$ and E_i is the energy of the i^{th} system state.

SIMPLE EXAMPLE – THE BABY ENSEMBLE
The following table illustrates the case of four systems A, B, C, and D with an available energy state spectrum consisting of with only four energies. For purposes of illustration, we assume nothing about the values of the E_i's, so while $N_T = 4$ we don't specify E_T except to say that it is a constant for the ensemble**.** We will further assume that our little ensemble is limited to exactly three possible distributions as seen in the first column of the following table. It should be noted however that not specifying the energies is equivalent to assuming that all four energies are the same in which case you can show that the actual number of possible ensemble states for our 'Baby Ensemble' is 256 = $(\text{number of states})^{N_T}$. For now, three distributions or 17 ensemble states is enough to get a feel for the process.

BABY ENSEMBLE

Ensemble States	**E1**	**E2**	**E3**	**E4**	Ω (n)	$P_{dist} = \frac{\Omega(n)}{\Omega_T(n)}$
Dist. 1	ABCD	-	-	-	1	1/17
Dist. 2	AB	C	D	-	12	12/17
	AB	D	C	-		
	AC	B	D	-		
	AC	D	B	-		
	AD	B	C	-		
	AD	C	B	-		
	BD	C	A	-		
	BD	A	C	-		
	BC	A	D	-		
	BC	D	A	-		
	CD	A	B	-		
	CD	B	A	-		
Dist. 3	ABC	-	-	D	4	4/17
	ABD	-	-	C		
	ACD	-	-	B		
	BCD	-	-	A		

Ensemble State:
An ensemble state is realized by specifying the number of systems, or occupation number of every energy state in the ensemble. However there is no way to know which systems are in a specific E_i … for even one system much less for all $N_T \to \infty$, so our only choice is to assume an *a-priori* approach where we are forced to assume that every system can visit every energy state consistent with our constraints on N_T and E_T.

Thus in our example above, it is not sufficient to say for example, that Dist. 2 has two systems in state E_1 and one system in each of the states E_2 and E_3 because these systems are macroscopic, can be labeled and therefore be realized in 12 equivalent ways. In this case we would say Dist. 2 represents 12 ensemble states or alternatively, Dist. 2 is 12 fold degenerate. Likewise, Dist.1 can be realized in one way and Dist. 3 in four ways. These numbers of systems in each distribution are called the 'occupation' numbers of the distributions, viz, if we write $(n) = \{n_1, n_2, ...\}$ then for Dist. 1: $n_1 = 4$, Dist. 2: $n_1 = 2$, $n_2 = 1$, $n_3 = 1$. For this simple case we can specifically compute the degeneracy of each distribution from the usual combinatorial expression $\Omega(n) = N_T! / \prod_i n_i!$ (see page 22 of your text) for distinguishable systems. This is how $\Omega(\text{n})$ in Column 6 of the table was calculated. The sum of the degeneracy's over all distributions is then the total degeneracy of the ensemble.

Thus our example ensemble has some 17 possible ensemble states based on three possible distributions, but real ensembles have nearly an infinite number of possible configurations. The question that arises is this: is there a preferred or most probable distribution amongst the huge number of possibilities? In our little ensemble, Dist. 2 is the most likely but for real systems it is impossible to know so we are forced to assume, a priori, the following:

Postulate 1):
All $\Omega_\text{T}(\text{ensemble}) = \sum_{\{\text{n}\}} \Omega(\text{n})$ ***ensemble states are postulated to be equally probable.***

Ensemble states are generated by specifying how many systems in the ensemble are in each of the (almost infinite) energy states E_1, E_2,.. available to every system and consistent with the constraints on E_T and N_T. Thus, in our example, we must assume all 17 possible states are equally probable.

STATEMENT OF THE PROBLEM:
Our basic goal is to compute P_i which can be thought of in following equivalent ways:

- Probability that a system chosen at random is in the i^{th} energy state E_i, or
- Fraction of the systems in the i^{th} energy state at any time, or
- Fraction of time that a system chosen at random is in the i^{th} energy state

Once we have P_i we can calculate averages of the mechanical variables like pressure and energy from the relationship: $\bar{G} = \sum_i P_i G_i$. This brings us to the second postulate:

Postulate 2):

It is assumed that the time average of any mechanical variable, G[p(t), q(t)]

$\bar{G}_{time} \equiv G_{obs} = \frac{1}{\tau}\int_{t_0}^{\tau+t_0} G[p(t),q(t)]dt$ ***where*** τ ***is large enough to eliminate fluctuations,***

is equal to the ensemble average of the same variable, i.e.,

$\bar{G}_{time} = \bar{G}_{ensemble} = \sum_{states} P_i G_i$ ***where the sum is over all system states.***

What we need then is an expression for calculating P_i for an arbitrary system chosen at random.

CALCULATION OF P_i:

Suppose our ensemble consists of a single distribution and for illustration purposes take Dist. 2 in the example above to represent our ensemble. Calculation of P_i, for all four energy states is straight forward, viz,

P_j for Distribution 2

	$P_i = n_i/N_T$
P_1	2/4
P_2	¼
P_3	¼
P_4	0

where $\Omega(n) = \Omega(n_1 = 2, n_2 = n_3 = 1, n_3 = 0)$. In column 2 we used the intuitive relationship for the probably, $P_i = n_i/N_T$, for a single distribution but this won't work for when we want P_i for an ensemble consisting of many distributions. In that case we must include the degeneracy for each distribution and then sum over all possible sets of occupation numbers. So our next task is to generalize $P_j = n_j/N_T$ to multiple distributions.

We write now for n_j

$$n_j = \left\{ \begin{matrix} \text{total systems in state j of distribution} \\ \text{n summed over all distributions} \end{matrix} \right\} = \sum_{\{n\}} \Omega(n) n_j(n)$$

Our probability is then,

$$P_j = \left\{ \begin{matrix} \text{fraction of systems in state} \\ \text{j for all distributions} \end{matrix} \right\} = \frac{n_j}{\{\text{total systems in all states of all distributions}\}}$$

$$= \frac{\sum_{\{n\}} \Omega(n) n_j(n)}{\sum_{states} \sum_{\{n\}} \Omega(n) n_j(n)}$$

Example: P_1 for the Baby Ensemble:

$$P_1 = \frac{[1*4 + 2*12 + 3*4]}{[1*4 + 2*12 + 3*4] + [0 + 1*12 + 0] + [0 + 1*12 + 0] + [0 + 0 + 1*4]} = 10/17$$

But, for any specific distribution n, we can write,

$$\sum_{states}\sum_{\{n\}}\Omega(n)n_j(n)=\sum_{states}n_j(n)\sum_{\{n\}}\Omega(n)=N_T\sum_{\{n\}}\Omega(n)$$

i.e., $N_T=\sum_{\{states\}}n_j(n)$ since for any particular distribution, the sum of the n_j over all possible energy states must add to the total number of systems in the ensemble, that is, every system in the ensemble must be accounted for when we sum over all states for a fixed distribution.

We finally have our desired recipe for calculating P_j, namely,

$$P_j=\frac{1}{N_T}\frac{\sum_{\{n\}}\Omega(n)\,n_j(n)}{\sum_{\{n\}}\Omega(n)}$$

and which is a completely general result for an arbitrary ensemble.

The sum in the numerator says: pick a j and then look at the first distribution. Multiply the number of systems in that distribution by its degeneracy. Then move onto the next distribution for that same j value. Do this for all distributions for a given j making sure to sum the values for each distribution. Then divide by the denominator which is just the total degeneracy times the number of systems. You then have P_j.

Applying our expression for P_j to our baby-ensemble above, we have:

P_1	[1/(4*17)] * [1*4 + 12*2 + 4*3] = 10/17
P_2	[1/(4*17)] * [12*1] = 3/17
P_3	[1/(4*17)] * [12*1] = 3/17
P_4	[1/(4*17)] * [4*1] = 1/17

Note that the P_i's sum to 1 as they should. We can also compute the average energy of the ensemble if we arbitrarily assign some values to the E_i's, eg, if E_1=5, E_2=10, E_3=15 and E_4=20 then $E_{avg}=\sum_i E_i P_i$ = 145/17 = 8.53.

We mentioned earlier that not specifying the values of the E_i's was tantamount to assuming they were all equal in which case one can show that the total degeneracy is given by

$$\Omega_T=\sum_{\{n\}}\frac{N_T\,!}{\prod_{n_i} n_i\,!}=[\text{number of states}]^{N_T}$$

and subject only to the restriction that

$$\sum_{\{n\}}n_i=N_T$$

Thus, our Baby Ensemble really has some $4^4 = 256$ possible energy states – something you might wish to confirm.

You can also apply our general expression for P_j to the previous case where we assumed the ensemble consisted of only distribution 2, to get P_j = [1/(4*12)] x [2*12] = ½ as before.

METHOD OF THE MOST PROBABLE DISTRIBUTION:

The expression for P_j that we have just derived is a general expression but as it stands is of limited value because the sums are not computable due to the fact that the number of systems is assumed to be infinite. That's the bad news. The good news is that we can take advantage of the bad news. *It is precisely the immensity of the problem that allows us to use the remarkable fact that for sums of this kind, there will always exist a single most probable distribution whose degeneracy is so overwhelmingly much greater than the degeneracy of all other distributions that our problem reduces to determining only one set of distribution numbers in our expression for $P_{i.}$ First though, we need to justify this extraordinary fact.*

PROOF OF THE MAXIMUM TERM METHOD

We now give a short proof of our assertion that 'only the largest term in the sum need be considered', is true. Following McQuarrie (also see Wikipedia – search on 'Maximum Term Method in Statistical Mechanics) we define S as the sum over the degeneracy, i.e.,

$S = \sum_{N=1}^{M} T_N = T_1 + T_2 + ... + T_M$ where T_N is positive for all N and the maximum value of T, i.e. $T_{max,}$ resides somewhere in the set T_1 to T_M but where is unknowable due to the immensity of the problem. For out purposes, it only matters that there is a maximum term. This enables us to write,

$$T_{max} \le S \le MT_{max}$$

which says that S is bounded by the maximum term, and the product of the number of terms M, and the value of the maximum term.

Taking logs of both sides changes nothing in terms of the relative bound so we have,

$$\ln T_{max} \le \ln S \le \ln M + \ln T_{max}$$

Now in statistical mechanics we find typically that $T_{max} \sim O(10^M)$ where $M \approx 10^{20}$ or larger, in which case we have,

$$10^{20} \le \log S \le 10^{20} + \log M$$

Thus, we find that $\ln S$ is bracketed by $\ln T_{max}$ since log M is negligible, that is,

$$\ln T_{max} \le \ln S \le \ln T_{max}$$

and conclude that $\ln S = \ln T_{max}$, that is, the sum itself is equal to its largest value, a result whose importance cannot be overstated.

Another way of viewing this result that emphasizes the nature of the factorial to grow really big in a hurry (in fact the only reason these approximations work at all is because of this property of the factorial) is the expression:

$\underset{M_{max}\to\infty}{\text{limit}} \log \sum_{k=1}^{M_{max}} k! \sim \log M!_{max}$ where, unlike the expression for S above, the sum is monotonically increasing. Note that again, we are replacing the sum by its largest term. So how good is this approximation? The error for M=10 is 0.7% and for M=150 it's already about 0.001%. Assuming the error for M $\approx 10^{20}$ or greater is vanishingly small we now derive the expression for the probability of the j^{th} system state P_j

$$P_j = \frac{1}{N_T} \frac{\sum_{\{n\}} \Omega(n)\, n_j(n)}{\sum_{\{n\}} \Omega(n)} \cong \frac{1}{N_T} \frac{\Omega(n^*)\, n_j(n^*)}{\Omega(n^*)} = \frac{n_j(n^*)}{N_T} \equiv \frac{n_j^*}{N_T}$$

where n_j^* is the number of systems in the j^{th} energy state ***of the most probable distribution,*** and that is precisely what we are after – provided of course that we can figure how to compute the value of n in the most probable distribution.

Our problem now boils down to determining the set of occupation numbers $\{n^*\}$ that correspond to the most probable distribution which is the one with the largest degeneracy. In other words we want to maximize,

$$\Omega(n) = \frac{N_T!}{\prod_i n_i!}$$

subject to the constraints,

$\sum_i n_i = N_T$ and,

$\sum_i n_i E_i = E_T$

Once we have the set {n*} we have $P_j = \frac{n_j^*}{N_T}$ which is the basis for everything we do from now on.

Maximizing $\Omega(n) = \frac{N_T!}{\prod_i n_i!}$ is done by taking the logs of Ω(n), using Stirlings approximation, adding in the constraints on N_T and E_T through the use of Lagrange Multipliers, differentiating and then solving for the general term n_i^* which maximizes Ω. This I leave to you folks (see problem 1-49 of your text).

Our result, after a bit of mucking about is:

$$n_j^* = N_T e^{-\alpha} e^{-\beta E_j}$$

which is the recipe for finding the set of occupation numbers that define the most probable distribution, $\Omega(n^*)$. The Lagrange Multipliers, α and β will be discussed in class and in your text, assuming you bought one.

DW McClure, Emeritus Professor of Chemistry/Physics
Portland State University
2/18/2010