

Types of ensembles

Marc R. Roussel
Department of Chemistry and Biochemistry
University of Lethbridge

March 6, 2009

So far, we have focused all our attention on the canonical ensemble, which is an ensemble of fixed, T , V and N . The canonical ensemble is a useful device for describing experiments in which the temperature is fixed by a heat bath. Of course, we can imagine other possibilities. Ensembles for some other common possibilities have in fact been devised. In this note, we discuss the other two ensembles commonly used to describe chemical systems, as well as their relationships to each other and to the canonical ensemble.

1 Some useful thermodynamic relationships

Recall that the chemical potential is the partial molar¹ Gibbs energy of a system, i.e.

$$dG = V dP - S dT + \mu dN$$

for a single-component system. Since $G = H - TS$,

$$dH = dG + T dS + S dT = V dP + T dS + \mu dN.$$

Similarly, since $H = U + PV$,

$$dU = dH - P dV - V dP = T dS - P dV + \mu dN.$$

Now if we isolate dS , we get

$$dS = \frac{1}{T} dU + \frac{P}{T} dV - \frac{\mu}{T} dN.$$

¹We're going to abuse both the terminology and the notation a little here. It is convenient to think in molecules rather than moles in much of what will follow. We therefore use the number of molecules N and express the chemical potential on a per molecule basis without changing our notation for this quantity.

Therefore,

$$T^{-1} = \left. \frac{\partial S}{\partial U} \right|_{V,N}. \quad (1)$$

$$\begin{aligned} P/T &= \left. \frac{\partial S}{\partial V} \right|_{U,N} \\ -\mu/T &= \left. \frac{\partial S}{\partial N} \right|_{U,V}. \end{aligned} \quad (2)$$

2 The microcanonical ensemble

An isolated system is both thermally insulated and mechanically undisturbed. According to the first law of thermodynamics, the internal energy of such a system is constant. We define the following ensemble to describe isolated systems:

A microcanonical ensemble is an ensemble of fixed U , V and N .

The basic postulate of statistical mechanics says that every microstate of an ensemble of isolated systems occurs with equal probability. Thus, if the ensemble has W different microstates, the probability that any particular member of the ensemble is in a particular microstate is

$$p_i = 1/W,$$

since the sum of all the probabilities have to add to 1 and they all must be equal. Another way to think about this equation is that the microcanonical partition function is just W .

The entropy of a microcanonical ensemble is calculated using the Boltzmann equation:

$$S = k \ln W. \quad (3)$$

Temperature can be defined for a microcanonical ensemble using equation 1. To apply this equation, we first note that the number of microstates available to the system depends on the energy, so $W = W(U)$. This means that S depends on U . We can therefore write

$$T = \left(\left. \frac{\partial S(U)}{\partial U} \right|_{V,N} \right)^{-1} = \left(k \left. \frac{\partial \ln W(U)}{\partial U} \right|_{V,N} \right)^{-1}. \quad (4)$$

Once we have T , we can define other quantities that depend on it, like the Helmholtz free energy:

$$A = U - TS = U - kT \ln W.$$

Example 2.1 Suppose that we have a microcanonical ensemble, each member of which consists of N harmonic oscillators of frequency ν holding a total of $U = Qh\nu$ units of energy, where Q is an integer and we have neglected zero-point energy. Earlier, we

found that the number of configurations of Q quanta of energy spread over N oscillators is

$$W = \frac{(N - 1 + Q)!}{(N - 1)!Q!}.$$

Suppose that N and Q are both large. This would be the case if, for instance, our oscillators were lattice vibrations in a simple solid and the temperature was not too low. Then

$$\begin{aligned} \ln W &= \ln(N - 1 + Q)! - \ln(N - 1)! - \ln Q! \\ &= (N - 1 + Q) \ln(N - 1 + Q) - (N - 1 + Q) \\ &\quad - [(N - 1) \ln(N - 1) - (N - 1)] - [Q \ln Q - Q] \\ &= (N - 1 + Q) \ln(N - 1 + Q) - (N - 1) \ln(N - 1) - Q \ln Q \\ &= (N + Q) \ln(N + Q) - N \ln N - Q \ln Q \end{aligned}$$

since we have already assumed that N and Q are large. We also have that $Q = U/(h\nu)$, so $\ln W$ depends on U . Thus,

$$\ln W = \left(N + \frac{U}{h\nu}\right) \ln \left(N + \frac{U}{h\nu}\right) - \frac{U}{h\nu} \ln \left(\frac{U}{h\nu}\right) - N \ln N.$$

Now applying equation 4, we get (after some algebra)

$$T = \left[\frac{k}{h\nu} \ln \left(\frac{Nh\nu}{U} + 1 \right) \right]^{-1} = \left[\frac{k}{h\nu} \ln \left(\frac{N}{Q} + 1 \right) \right]^{-1}.$$

Let us consider a couple of extremes:

1. As $N/Q \rightarrow \infty$, $T \rightarrow 0$. This would be the case where there were few quanta stored relative to the number of oscillators.
2. As $N/Q \rightarrow 0$, $T \rightarrow \infty$. This limit corresponds to the case where there are many quanta for each oscillator.

Intuitively, we expect the temperature to ramp up with the amount of energy stored (i.e. with the value of Q). Both of these results are consistent with this intuition.

3 The grand canonical ensemble

Both the microcanonical and canonical ensembles deal with systems with a fixed number of particles N (or, in general, fixed values of the numbers of particles of each type, N_1, N_2, \dots). However, the number of molecules of a given type in a chemical system at equilibrium is not generally fixed.² Due to chemical reactions, the number of molecules of each type fluctuates with time. These fluctuations are generally fairly small, but they are not unimportant from the point of view of understanding how chemical systems behave. We address this problem by fixing the chemical potential instead of the number of particles:

²We won't even touch the much more delicate issue of systems which are out of equilibrium.

A grand canonical ensemble is an ensemble of fixed T , V and μ .

Imagine that we put two systems of constant volume together in such a way that they can exchange energy and particles. This could be as simple as opening a door between two rigid compartments. The two compartments are not necessarily the same size. In fact, we will eventually let one become much larger than the other. The composite system is isolated from the outside world, i.e. it is a sample from a microcanonical ensemble. We now ask what equilibrium conditions hold in such a system. Suppose that the energies of the two subsystems are E_1 and E_2 , and that the numbers of molecules in each subsystem are N_1 and N_2 . Since the system is isolated, the total energy and number of molecules are conserved:

$$\begin{aligned} U_1 + U_2 &= U_T, \\ \text{and } N_1 + N_2 &= N_T, \end{aligned}$$

where U_T and N_T are constants. On the other hand, because energy and particles flow freely between the two systems, U_1 , U_2 , N_1 and N_2 all fluctuate with time. The number of microstates in a microcanonical ensemble depends in general on U , N and V . In this particular case, we are less interested in the dependence on V since the volumes of the two compartments are fixed, unlike U_i and N_i . The total number of microstates of the composite system is the product of the number of microstates of the two subsystems:

$$W(U_T, N_T) = W_1(U_1, N_1)W_2(U_2, N_2) = W_1(U_1, N_1)W_2(U_T - U_1, N_T - N_1).$$

The number of microstates in each subsystem will depend on how the energy and particles are shared between the two compartments, so the overall number of microstates depends on U_1 and N_1 for fixed values of U_T and N_T . As we have seen previously, there will however be a dominant configuration which is overwhelmingly more probable than any other. This is the configuration that maximizes $W(U_T, N_T)$. For the composite microcanonical system, the entropy is a strictly increasing function of W (equation 3), so maximizing W is the same as maximizing S . According to the second law of thermodynamics, the state of maximum entropy is of course the equilibrium state of the system. Entropy is additive, so

$$S = S_1(U_1, N_1) + S_2(U_T - U_1, N_T - N_1).$$

To find the equilibrium conditions, differentiate this equation with respect to both U_1 and N_1 and set these derivatives equal to zero:

$$\begin{aligned} \left. \frac{\partial S_1}{\partial U_1} \right|_{V, N_1} + \left. \frac{\partial S_2}{\partial U_1} \right|_{V, N_1} &= \left. \frac{\partial S_1}{\partial U_1} \right|_{V, N_1} - \left. \frac{\partial S_2}{\partial U_2} \right|_{V, N_1} = 0, \\ \left. \frac{\partial S_1}{\partial N_1} \right|_{V, U_1} + \left. \frac{\partial S_2}{\partial N_1} \right|_{V, U_1} &= \left. \frac{\partial S_1}{\partial N_1} \right|_{V, U_1} - \left. \frac{\partial S_2}{\partial N_2} \right|_{V, U_1} = 0. \end{aligned}$$

To obtain the second equality in each of these two equations, we applied the chain rule to $S_2(U_2, N_2) = S_2(U_T - U_1, N_T - N_1)$. Now we apply equations 1 and 2 and rearrange slightly:

$$\begin{aligned} \frac{1}{T_1} &= \frac{1}{T_2}, \\ \frac{\mu_1}{T_1} &= \frac{\mu_2}{T_2}. \end{aligned}$$

The first equation tells us that the temperatures of the two subsystems will be the same at equilibrium, perhaps not surprisingly. The second equation combined with the first tells us that their chemical potentials will also be the same. Thus, this construction gives us two members of a grand canonical ensemble.

Now that we know that our two subsystems belong to a grand canonical ensemble, we would like to determine the probability of obtaining one of these states, which in turn will lead us to the grand canonical partition function. Let us consider the entropy of subsystem 2:

$$S_2(U_T - U_1, N_T - N_1) = k \ln W_2(U_T - U_1, N_T - N_1),$$

$$\text{or } W_2(U_T - U_1, N_T - N_1) = e^{S_2(U_T - U_1, N_T - N_1)/k}.$$

Suppose that subsystem 2 is much larger than subsystem 1, i.e. that $U_1 \ll U_T$ and $N_1 \ll N_T$. Then, by Taylor's theorem,

$$S_2(U_T - U_1, N_T - N_1) \approx S_2(U_T, N_T) - U_1 \left. \frac{\partial S_2}{\partial U} \right|_{V, N_T} - N_1 \left. \frac{\partial S_2}{\partial N} \right|_{U_T, V} = S_2(U_T, N_T) - \frac{U_1}{T} + \frac{N_1 \mu}{T},$$

again using equations 1 and 2. The number of microstates of system 2 is therefore

$$W_2(U_T - U_1, N_T - N_1) = e^{S_2(U_T, N_T)} e^{-(U_1 - N_1 \mu)/(kT)} = W_2(U_T, N_T) e^{-(U_1 - N_1 \mu)/(kT)}.$$

The number of microstates should be proportional to the probability of subsystem 1 having U_1 units of energy and N_1 particles. The only term in this equation which depends on U_1 and N_1 is the last one, so the probability of having U_1 units of energy and N_1 particles in the grand canonical ensemble must be proportional to this term. Thus we have

$$p(U, N) = \frac{e^{-(U - N\mu)/(kT)}}{Q_G}, \quad (5)$$

dropping the subsystem subscript, where Q_G is **grand canonical partition function**. This is obtained by summing the probability over all possible values of U and N . Since the exponential factors into $e^{-U/(kT)} e^{N\mu/(kT)}$, the sum over all possible values of U just gives a factor of the canonical partition function, Q . The result is

$$Q_G(\mu, T, V) = \sum_{N=0}^{\infty} e^{N\mu/(kT)} Q(T, V, N). \quad (6)$$

The probability that a particular member of a grand canonical ensemble has energy U and contains exactly N particles at a particular instant is given by equation 5. We can now calculate the statistics of the variables U and N in an open system that is in equilibrium with its thermal and chemical surroundings. In particular, we might be interested in the size of the fluctuations in U or N . Here we tackle the fluctuations in N . As an exercise, I suggest that you try to derive an equation for the size of the fluctuations in U .

The standard deviation is an obvious measure of the size of the fluctuations in a variable. Recall that it is calculated by

$$\sigma_N^2 = \langle N^2 \rangle - \langle N \rangle^2,$$

taking the statistics of N as an example. The averages are calculated as follows:

$$\begin{aligned}
\langle N \rangle &= \sum_U \sum_{N=1}^{\infty} N p(U, N) \\
&= \sum_U \sum_{N=1}^{\infty} N e^{-(U-N\mu)/(kT)} / Q_G \\
&= \frac{1}{Q_G} \sum_U \sum_{N=1}^{\infty} N e^{-(U-N\mu)/(kT)} \\
&= \frac{1}{Q_G} \sum_{N=1}^{\infty} N e^{N\mu/(kT)} Q(T, V, N). \\
\therefore Q_G \langle N \rangle &= \sum_{N=1}^{\infty} N e^{N\mu/(kT)} Q(T, V, N).
\end{aligned}$$

Now we use a trick. (There is no reason you should have seen this coming.) Differentiate this equation with respect to μ :

$$\begin{aligned}
\langle N \rangle \sum_{N=0}^{\infty} \frac{N}{kT} e^{N\mu/(kT)} Q(T, V, N) + Q_G \left. \frac{\partial \langle N \rangle}{\partial \mu} \right|_{T,V} &= \sum_{N=0}^{\infty} \frac{N^2}{kT} e^{N\mu/(kT)} Q(T, V, N). \\
\therefore \langle N \rangle \frac{1}{Q_G} \sum_{N=0}^{\infty} N e^{N\mu/(kT)} Q(T, V, N) + kT \left. \frac{\partial \langle N \rangle}{\partial \mu} \right|_{T,V} &= \frac{1}{Q_G} \sum_{N=0}^{\infty} N^2 e^{N\mu/(kT)} Q(T, V, N). \\
\therefore \langle N \rangle^2 + kT \left. \frac{\partial \langle N \rangle}{\partial \mu} \right|_{T,V} &= \langle N^2 \rangle. \tag{7}
\end{aligned}$$

To get this last relationship, we used

$$\langle N^2 \rangle = \frac{1}{Q_G} \sum_{N=1}^{\infty} N^2 e^{N\mu/(kT)} Q(T, V, N),$$

as well as some intermediate results from above. We can rearrange equation 7 to

$$\sigma_N^2 = \langle N^2 \rangle - \langle N \rangle^2 = kT \left. \frac{\partial \langle N \rangle}{\partial \mu} \right|_{T,V}. \tag{8}$$

Recall that, for a single-component system, $G = N\mu$. If the number of molecules is not fixed, then the average free energy should be $\langle G \rangle = \langle N \rangle \mu$. If we take a differential of this expression, we get $d\langle G \rangle = \mu d\langle N \rangle + \langle N \rangle d\mu$. We also have $d\langle G \rangle = -S dT + V dP + \mu d\langle N \rangle$. If we set these two equations for $d\langle G \rangle$ equal to each other, we get $\langle N \rangle d\mu = -S dT + V dP$. At constant temperature, this relationship becomes $d\mu = (V/\langle N \rangle) dP$, but $V/\langle N \rangle = V_m$, the molar volume,³ so $d\mu = V_m dP$. This equation holds when we change these variables

³Again, I'm abusing the terminology a little here since this is the volume per molecule, not per mole.

by changing any independent at constant temperature. In particular, we could change the molar volume:

$$\left. \frac{\partial \mu}{\partial V_m} \right|_T = V_m \left. \frac{\partial P}{\partial V_m} \right|_T. \quad (9)$$

This equation will hold regardless of how we change the molar volume. We can change the molar volume either by changing $\langle N \rangle$ at constant V or V at constant $\langle N \rangle$. We do the former on the left-hand side, and the latter on the right-hand side:

$$\begin{aligned} \left. \frac{\partial \mu}{\partial V_m} \right|_{T,V} &= \left. \frac{\partial \mu}{\partial \langle N \rangle} \right|_{T,V} \left. \frac{\partial \langle N \rangle}{\partial V_m} \right|_{T,V} \\ \left. \frac{\partial V_m}{\partial \langle N \rangle} \right|_{T,V} &= -\frac{V}{\langle N \rangle^2} \\ \therefore \left. \frac{\partial \mu}{\partial V_m} \right|_{T,V} &= -\frac{\langle N \rangle^2}{V} \left. \frac{\partial \mu}{\partial \langle N \rangle} \right|_{T,V} \\ \left. \frac{\partial P}{\partial V_m} \right|_{\langle N \rangle, T} &= \left. \frac{\partial P}{\partial V} \right|_{\langle N \rangle, T} \left. \frac{\partial V}{\partial V_m} \right|_{\langle N \rangle, T} = \langle N \rangle \left. \frac{\partial P}{\partial V} \right|_{\langle N \rangle, T}. \end{aligned}$$

Now going back to equation 9, we get

$$\begin{aligned} -\frac{\langle N \rangle^2}{V} \left. \frac{\partial \mu}{\partial \langle N \rangle} \right|_{T,V} &= V \left. \frac{\partial P}{\partial V} \right|_{\langle N \rangle, T} \\ \therefore \left. \frac{\partial \langle N \rangle}{\partial \mu} \right|_{T,V} &= -\frac{\langle N \rangle^2}{V^2} \left. \frac{\partial V}{\partial P} \right|_{\langle N \rangle, T}. \end{aligned}$$

However, the isothermal compressibility is defined by

$$\begin{aligned} \kappa &= -\frac{1}{V} \left. \frac{\partial V}{\partial P} \right|_{N,T} \\ \therefore \left. \frac{\partial \langle N \rangle}{\partial \mu} \right|_{T,V} &= \frac{\kappa \langle N \rangle^2}{V}. \end{aligned}$$

We can now go back to equation 8:

$$\begin{aligned} \sigma_N^2 &= \frac{kT\kappa \langle N \rangle^2}{V} \\ \therefore \frac{\sigma_N}{\langle N \rangle} &= \sqrt{\frac{kT\kappa}{V}}. \end{aligned}$$

This is the result we were looking for. From this equation, we can predict the relative size of fluctuations in particle numbers in an open system, for instance in a particular volume of a gas.

Example 3.1 For an ideal gas, $\kappa = 1/P$, so

$$\frac{\sigma_N}{\langle N \rangle} = \sqrt{\frac{kT}{PV}} = \sqrt{\frac{1}{N}}.$$

If we have a volume large enough to hold (on average) 1 mol of gas, then we expect relative fluctuations in the number of molecules in any given measurement of the order of $1/\sqrt{N_A} \approx 10^{-12}$. These would be undetectable. On the other hand, if we are studying a reaction in a small volume of a gas at a very low density, we might begin to see fluctuations.

4 How the ensembles are related

The names “microcanonical”, “canonical” and “grand canonical” suggest a hierarchy of ensembles. This is intentional.

Suppose that we have a very large grand ensemble consisting of a number of identical volumes which communicate with each other. Imagine that we simultaneously shut all the doors connecting these volumes. The number of molecules in each volume is now fixed. We can extract from the original grand ensemble a set of canonical ensembles by collecting together all the volumes which have the same value of N . In a sense then, a grand ensemble is an ensemble of canonical ensembles. We can see this relationship in the form of the grand partition function (equation 6): It is an average of canonical partition functions weighted by a function of N .

Similarly, suppose that we have a very large canonical ensemble consisting of a number of identical replicas in a thermostat device at temperature T . If we simultaneously isolate all the replicas, we now have a number of systems each of which has a fixed value of U . We can make microcanonical ensembles by collecting together all the members of the original canonical ensemble with the same value of U .

The description of each ensemble tells us what application it is intended for. However, we often cheat and use whatever ensemble is convenient in calculations, regardless of the actual system we are trying to study. Why can we do that? The reason can be glimpsed in example 3.1: Fluctuations from the ensemble average are typically very small. This applies to any quantity which is variable in an ensemble, so for example the fluctuations in temperature in a microcanonical ensemble, even though they are allowed since T is not held constant in this ensemble, are very small. It therefore makes very little difference in practice whether we hold U or T constant, or N or μ , unless the system we are interested in is either very small or has some mechanism for amplifying fluctuations.