Homework 3: Canonical ensemble and radial distribution function

Glen Hocky

Due: October 13, 2020

1. (Computational) This exercise will demonstrate how to compute a radial distribution function ($g(r)$) for somewhat realistic simulation data.
   It is available by logging in to https://chemga-2600-fall.rcnyu.org/ and going to the assignments tab. There you can push "Fetch" next to RDF. From here you can access the RadialDistributionFunction exercise. When done, click validate and then submit.

   Alternatively, fill in the missing pieces in the notebook within the rdf folder on the course github:
   https://github.com/hockyg/chem-ga-2600/

   When finished, download with File->Download As->ipynb so you have a copy and send it in by email or slack.

2. *Ideal gas of molecules* (Adapted from Tuckerman problem 4.12). If you have N identical non-interacting molecules (each with n atoms) in a box, the total partition function factorizes,
$$Q(N, V, T) = \frac{q(n, V, T)^N}{N!} \tag{1}$$
   Moreover, the single molecule partition function $q(n, V, T) = V f(n, T)$, where f is a function that only depends on the number of atoms and the temperature.
   Now suppose that the system contains different types of molecules, this factorization still works. For example, with 2 types, A and B,
$$Q(N_A, N_B, V, T) = \frac{q_A(n_A, V, T)^{N_A}}{N_A!} \cdot \frac{q_B(n_B, V, T)^{N_B}}{N_B!} \tag{2}$$

   (a) An example chemical reaction might be
$$aA + bB \rightleftharpoons cC + dD \tag{3}$$

The Helmholtz free energy A is now a function of V,T, and all 4 N's. Let there be a variable called λ which is the reaction extent. Then $dN_A = a\,d\lambda$, $dN_B = b\,d\lambda$, $dN_C = -c\,d\lambda$, and $dN_D = -d\,d\lambda$. At chemical equilibrium, A is at a minimum, which means that $dA/d\lambda = 0$.

Show that at equilibrium with fixed V and T,

$$a\mu_A + b\mu_B - c\mu_C - d\mu_D = 0, \tag{4}$$

where chemical potentials of each species are defined like:

$$\mu_A = -k_B T \frac{\partial \ln Q(V,T,N_A,N_B,N_C,N_D)}{\partial N_A} \tag{5}$$

Hint: This should follow very directly from writing out the chain rule for dA.

(b) By plugging in for Q in the 4 equations like Eq. 5 and substituting this in to Eq. 4, show that you get the following relationship

$$K(T) = \frac{\rho_C^c \rho_D^d}{\rho_A^a \rho_B^b} = \frac{(q_C/V)^c (q_D/V)^d}{(q_A/V)^a (q_B/V)^b}, \tag{6}$$

where $\rho_A = N_A/V$, etc. Hint, you may have to use Sterling's approximation. Do you see why both the middle and right fractions are functions of only temperature?

(c) Using the definition of the Helmholtz free energy and the formula for Q with 4 different species, show that the Helmholtz free energy A can be written as $A_{total} = A_A + A_B + A_C + A_D$. What is the equation for A_X, that is, the contribution to the Helmholtz free energy from species X.

(d) The pressure in the canonical ensemble can be obtained by the formula $P = -\left(\frac{\partial A}{\partial V}\right)_{N,T}$. The partial pressure of species X can be expressed similarly as $P_X = -\left(\frac{\partial A_X}{\partial V}\right)_{N,T}$, where A_X was defined in the previous part. Given this definition, find the relationship between K(T) above and

$$K_P = \frac{P_C^c P_D^d}{P_A^a P_B^b} \tag{7}$$

3. *Deriving the Canonical Ensemble from the perspective of Maximum Entropy.* Consider a system with N possible states each having energy ϵ_i. One definition we read for the entropy is the Gibbs entropy, $S = -k_B \sum_{i=1}^N p_i \ln p_i$, where p_i is the (unknown) probability that system is in state p_i.

By definition, $\sum_{i=1}^N p_i = 1$, and the average energy $\bar{E}$ is a constant, with $\bar{E} = \sum_{i=1}^N \epsilon_i p_i$. Note that the Gibbs entropy is strictly positive since $0 < p_i < 1 \Rightarrow -\ln p_i > 0$.

A question we could ask is, which set of $\{p_i\}$ maximize the Gibbs entropy subject to the constraints $\sum_{i=1}^{N} p_i = 1$ and $\bar{E} = \sum_{i=1}^{N} \epsilon_i p_i$. To do this, we can use the method of undetermined Lagrange Multipliers. If you don't know what this is, you should read about it a little on your own. The upshot is, we have to maximize the quantity:

$$L = -k_B \sum_{i=1}^{N} p_i \ln p_i - \alpha \sum_{i=1}^{N} p_i - \beta \sum_{i=1}^{N} \epsilon_i p_i, \tag{8}$$

where α and β are undetermined variables.

(a) Find where L is maximum as a function of an arbitrary p_k by solving $\frac{dL}{dp_k} = 0$ for p_k, where the p's are considered all independent variables.

(b) Use the fact that $\sum_{i=1}^{N} p_i = 1$ to eliminate one of the unknown variables. Your expression for the probabilities should look familiar.

(c) Based on what you know about the Canonical ensemble, express this β in terms of a familiar thermodynamic variable.

(d) Optional extra part: what happens if we don't include the constraint on the average energy? What ensemble does this correspond to?