Date: Thursday, November 9

Instructions: You may utilize notes and problem set solutions (both your solutions and the posted solutions). You may not, however, discuss the problems with others.

Problem 1:	/ 12
Problem 2:	/ 15
Problem 3:	/ 16
Problem 4:	/ 30
Problem 5:	/ 27
Total:	/ 100

Equations you may find useful:

$$1 = \int_{-\infty}^{\infty} dx\, \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$\mu = \int_{-\infty}^{\infty} dx\, \frac{x}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$\sigma^2 = \int_{-\infty}^{\infty} dx\, \frac{(x-\mu)^2}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$P(\nu) = \frac{e^{-\beta E(\nu)}}{Q(\beta)} \quad \text{[Canonical]}$$

$$S = k_{\mathrm{B}} \ln \Omega$$

$$-\beta A = \ln Q \quad \text{[Canonical]}$$

$$\beta = \frac{1}{k_{\mathrm{B}} T} = \frac{1}{k_{\mathrm{B}}} \left(\frac{\partial S}{\partial E}\right)_{N,V}$$

$$C_V = \left(\frac{\partial \langle E \rangle}{\partial T}\right)_{N,V}$$

$$Q(\beta) = \sum_{\nu} e^{-\beta E(\nu)} \quad \text{[Canonical]}$$

$$\ln n! \approx n \ln n - n$$

$${}_{M}C_{N} = \frac{M!}{N!(M-N)!}$$

1. **System size scaling. [12 pts.]** For each of the following, identify the dependence on N. Answers should be in the form of a proportionality. For example, you might answer that the object is proportional to N, proportional to $\ln N$, proportional to 1, or proportional to some other function of N. (Saying something is proportional to 1 is another way of saying there is no N dependence.) For full credit, also provide a brief rationale for each answer.

   (i) The number of classical microstates for N particles to be arranged in a box of size V with energy E: $\Omega(N, V, E)$. **[2 pts.]**

   [Hint: You may want to subdivide the system into M independent cells, each with volume v, density $\rho = N/V$, and energy density $\epsilon = E/N$. Let the number of microstates of one such cell be $\tilde{\omega}$. Your answer will involve $\omega = \tilde{\omega}^{1/\rho v}$.]

   $\Omega(N, V, E) \propto e^{N \ln \omega}$. The combined system formed from the M cells must have $\tilde{\omega}^M = \left(\tilde{\omega}^{1/\rho v}\right)^N$ microstates. Hence $\Omega \propto \omega^N$. Another way of writing that is $\Omega \propto e^{N \ln \omega}$.

   (ii) The entropy of a material with N particles in a volume V with energy E: $S(N, V, E)$. **[2 pts.]**

   $S(N, V, E) \propto N$. Since $S = k_{\mathrm{B}} \ln \Omega$, this follows directly from the exponential scaling of Ω.

   (iii) The Gibbs free energy of a material with N particles kept at pressure p and temperature T: $G(N, p, T)$. **[2 pts.]**

   $G(N, p, T) \propto N$. Recall that $G = E - TS + pV$. On the right-hand side, E, S, and V are all extensive while T and p are intensive. It follows that the sum is extensive.

   (iv) The inverse temperature of a N-particle bath of volume V: $\beta = \frac{1}{k_{\mathrm{B}}}\left(\frac{\partial S}{\partial E}\right)_{N,V}$. **[2 pts.]**

   $\beta \propto 1$. Notice that both S and E are extensive, so the derivative of one with respect to the other has no N dependence.

   (v) The mean squared length between endpoints of a **one**-dimensional lattice polymer: $\langle R^2 \rangle$. As you hopefully recall from homework, each bond of the lattice polymer is equally likely in all directions, irrespective of the other bonds. For a one dimensional polymer that means steps left and right, each occur with probability $1/2$. **[2 pts.]**

   $\langle R^2 \rangle \propto N$. The key fact is that the bonds choose their directions independently. Letting $\vec{R}_i$ be the displacement vector for bond i, we have a net displacement of

   $$\vec{R} = \sum_{i=1}^{N} \vec{R}_i.$$

   When squaring that displacement, it is useful to write R^2 as a product of a sum over dummy variable i times another sum over dummy variable j. We then separate the cross-terms of that multiplication into terms with $i = j$ and those with $i \neq j$:

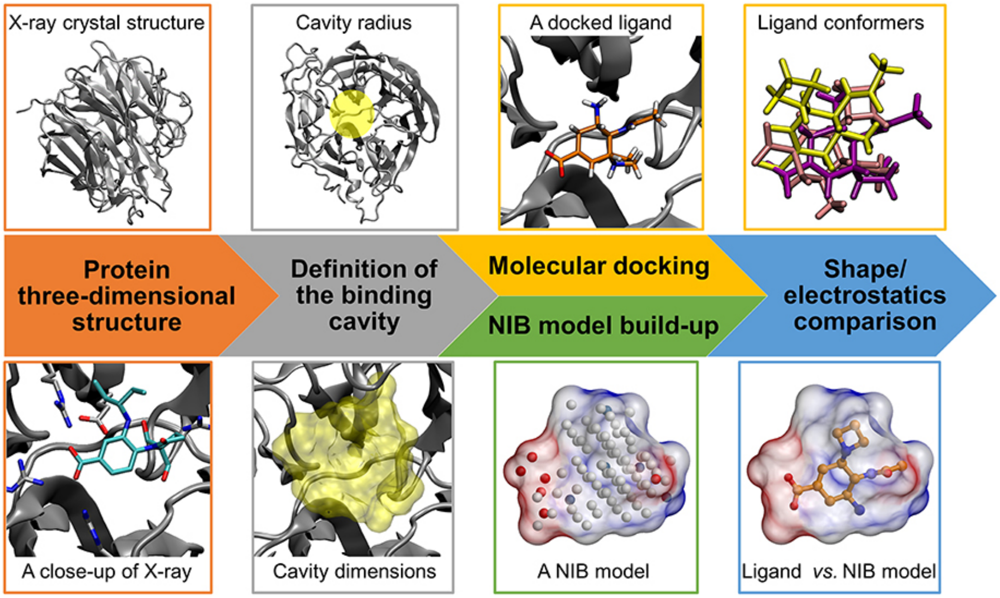
   $$R^2 = \sum_i R_i^2 + \sum_i \sum_{j \neq i} \vec{R}_i \cdot \vec{R}_j.$$

The terms in the second sum average to zero since bond i is independent of bond j when $i \neq j$. Hence $\left\langle R^2 \right\rangle = N \left\langle R_1^2 \right\rangle \propto N$.

(vi) The mean squared length between endpoints of a **three**-dimensional lattice polymer: $\left\langle R^2 \right\rangle$. Each bond is still independent of the others and each of the six directions is still equally likely. **[2 pts.]**

$\left\langle R^2 \right\rangle \propto N$. The key fact is that the bonds choose their directions independently, so the same logic as in (v) will hold.

2. **Protein-ligand binding. [15 pts.]** The pharmaceutical industry routinely wrestles with the challenge of identifying small molecules which can act as drugs by binding to disease-related proteins. For decades they have used computers to help screen through candidates of potential drugs. A common approach to such screens, known as docking, is summarized in the following picture:



Briefly, the idea is to take the structure of a frozen protein from a crystal, understand the shape and charge distribution around some binding pocket (cavity), then computationally try to fit different ligands into that pocket. The picture refers to that last step as “shape/electrostatics comparison”.

You can perhaps imagine that powerful quantum chemistry techniques like density functional theory could be used to compute the energy for the protein configuration in the presence of the ligand, and you could seek ligands which minimize that energy. That approach makes me a little uncomfortable because it has “integrated out” all of the solvent degrees of freedom. My stat mech upbringing makes me worried that ignoring the solvent could cause problems.

To make me feel more comfortable, Roel proposes that he could offer me a computer big enough to explicitly include the solvent. He says that if I can provide him with a microstate (a configuration of the protein, ligand, and surrounding solvent molecules), he could compute for me the energy of that microstate. He suggests I could identify ligands that bind tightly to the protein by looking for ligands with a low microstate energy.

Explain why computing those microstate energies does not fully indicate whether a drug will be likely to bind to the cavity. A good answer will *discuss* the issue(s) in a paragraph or two. You may even be able to indicate how Roel’s capabilities could be useful as a piece of a more complete approach. Feel free to use analogies, plots, drawings, or any other tools that help you make your point(s) more clearly.

The main thing I wanted you to address was that “whether a drug will be likely to bind to the cavity” is fundamentally a question about the probability of a macrostate (bound macrostate versus unbound macrostate). One cannot determine the probability of a macrostate using the probability of a single microstate. So, no matter how good the magical quantum mechanical calculation is at giving a “true” energy for the protein + ligand + solvent configuration, you would only be getting the Boltzmann probability for that one configuration. The probability for the bound macrostate requires that you know the total probability that comes from summing the Boltzmann probability for *all* bound microstates and comparing that sum to the total Boltzmann probability for *all* unbound microstates. The combination of energetics (feeding into the $e^{-\beta E}$ of the Boltzmann weights) and entropics (having to do with how many microstates there are in the bound and unbound sets) can be recast as a *free energy* of binding. The quantum mechanical calculation might be very useful at helping you accurately get the energetic part. That would mean that if I gave you two configurations, x and y, you would be able to give me a very accurate estimate for how much more or less probable it is to find configuration x versus configuration y at temperature T: $p(x)/p(y) = e^{-(E(x)-E(y))/k_\mathrm{B}T}$. That can be a useful piece of computing a free energy, but you would need to also *sample* different possible microstate configurations as a way of measuring the difference in the entropy of the two macrostates. Answers varied widely, so Todd graded these responses holistically.

3. **Fluctuating energy and particles. [16 pts., 2 pt. each]** Imagine a system surrounded by a rigid, permeable wall. This system can exchange both energy and particles with a much larger reservoir which has an inverse temperature β and a chemical potential μ. Fill in the following blanks in the description of this ensemble.

   The equilibrium probability for microstate ν is

   $$P(\nu) = \begin{cases} \dfrac{\exp\left(-\beta E(\nu) + \boxed{\beta\mu N(\nu)}\right)}{\Xi\left(\boxed{\mu, V, T}\right)}, & \text{if } \boxed{V(\nu) = V} \\ 0, & \text{otherwise,} \end{cases}$$

   where

   $$\Xi = \sum_{\nu \text{ with } V(\nu)=V} \exp\left(\boxed{-\beta E(\nu) + \beta\mu N(\nu)}\right).$$

   Various Legendre transforms of the internal energy yield thermodynamic potentials. A list of such transformations includes: $E, A = E - TS, G = E - TS + pV, H = E + pV, \mathcal{F} = E - \mu N, \Phi = E - TS - \mu N, W = E + pV - \mu N$. The partition function Ξ can be connected to one of these thermodynamic potentials as

   $$-k_{\mathrm{B}} T \ln \Xi = \boxed{\Phi}.$$

   Let us call that thermodynamic potential ☹ (so as not to give away the previous answer). A small change in ☹ could be related to small changes in the three natural variables as

   $$d☹ = \boxed{-SdT - pdV - Nd\mu}$$

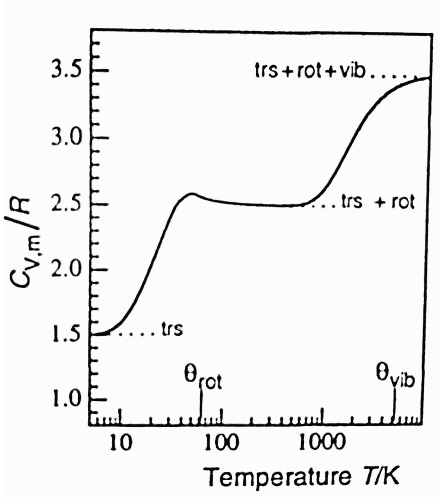
   The partition function can furthermore be used as a generating function. The first two derivatives yield

   $$\left(\frac{\partial \ln \Xi}{\partial(\beta\mu)}\right)_{T,V} = \boxed{\langle N \rangle}$$

   and

   $$\left(\frac{\partial^2 \ln \Xi}{\partial(\beta\mu)^2}\right)_{T,V} = \boxed{\langle \delta N^2 \rangle}.$$

4. **Heat capacity. [32 pts.]** The heat capacity per mole of a diatomic gas, HD, is plotted below in terms of the gas constant $R = N_\mathrm{A} k_\mathrm{B}$, with N_A being Avogadro's number.



Based on the labels and your chemical intuition, you see that as you move to higher temperatures you "unlock" higher energy states. At low temperatures the molecules have translational motion but are stuck in the ground rotational and vibrational states. At a high enough energy, the rotational excitations start to become relevant, and at a still higher temperature the vibrations also influence the heat capacity.

Todd tries to claim that the HD heat capacity should be simple to understand—the heat capacity should go up the more microstates you unlock. You point out to him that the behavior around $T = \theta_\mathrm{rot}$ seems to contradict his simplistic explanation.

The population of the excited states will monotonically increase as temperature increases, but the heat capacity has a peak! This problem will help you convince Todd that a non-monotonic heat capacity is not a concern (and indeed could have been anticipated quite simply).

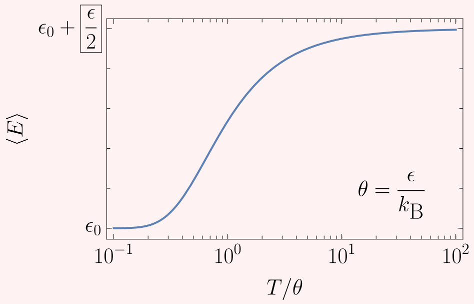
(i) The two-state model you studied in Problem Set 6 is a good starting point. Remember, that model has only two possible energy levels that differ in energy by ϵ. Demonstrate that the probability of the excited state in the two-state model increases monotonically with temperature. **[5 pts.]**

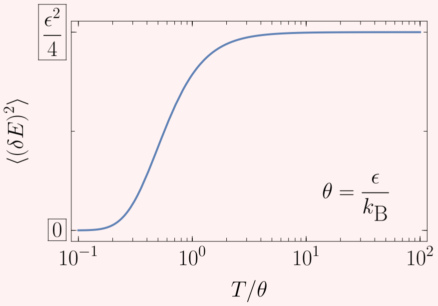
$$P_\uparrow = \frac{e^{-\beta\epsilon}}{1+e^{-\beta\epsilon}} = (1+e^{\beta\epsilon})^{-1} \quad \Longleftrightarrow \quad \frac{\mathrm{d}P_\uparrow}{\mathrm{d}\beta} = -\epsilon(1+e^{\beta\epsilon})^{-2} < 0.$$

Because T increases as β decreases, this implies that $P_\uparrow$ increases monotonically with T.

(ii) *Roughly sketch* plots of average energy $\langle E \rangle$ and of fluctuations in energy $\langle \delta E^2 \rangle$ for the two-state model. Fill in the boxes to label the axes with the correct limiting behavior. In case it is not obvious, when I ask for a rough sketch, I'm paying attention to a few distinguishing features like: does it go up, does it go down, what are the low-temperature limits, what are the high-temperature limits, how many peaks does it have, where are the peaks, etc. **[8 pts.]**

As $T \to 0$, we know that the system must solely be in the low-energy state with energy ϵ_0, so $\langle E \rangle = \epsilon_0$. Similarly, there won't be fluctuations from this energy (because all particles will be in the low-energy state), so $\langle \delta E^2 \rangle = 0$. As $T \to \infty$, we know that the system exists in both states with equal probability, so $\langle E \rangle = \epsilon_0 + \epsilon/2$. We also have $\langle \delta E^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2 = \epsilon^2/2 - \epsilon^2/4 = \epsilon^2/4$. Because it wouldn't make sense for the variance to depend on ϵ_0—translating a probability distribution shouldn't affect its variance—we can choose $\epsilon_0 = 0$ to simplify our calculation.





The important points we checked were the three boxed labels, that each curve crossed over from one plateau at low T to the high-temperature plateau, and that the cross-over occurred around $T = \theta$.

(iii) Derive the two-state model's heat capacity:

$$C_V = \frac{\epsilon^2 e^{-\epsilon/(k_B T)}}{k_B T^2 \left(1 + e^{-\epsilon/(k_B T)}\right)^2}.$$

Don't be shy about including some sentences that explain what you are doing in each step. **[5 pts.]**

The heat capacity can be obtained by taking two derivatives of $Q(\beta)$ with respect to β:

$$C_V = \frac{\partial \langle E \rangle}{\partial T} = -\frac{\partial \beta}{\partial T} \left(\frac{\partial^2 \ln Q}{\partial \beta^2} \right)_{N,V} = \frac{1}{k_B T^2} \left(\frac{\partial^2 \ln Q}{\partial \beta^2} \right)_{N,V}.$$

For the two-state model, Q is quite forward to compute and differentiate. In terms of the ground state energy ϵ_0,

$$Q = e^{-\beta \epsilon_0} + e^{-\beta(\epsilon_0 + \epsilon)} = e^{-\beta \epsilon_0} \left(1 + e^{-\beta \epsilon}\right).$$

The first derivative gives

$$\left(\frac{\partial \ln Q}{\partial \beta} \right)_{N,V} = -\epsilon_0 - \frac{\epsilon e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}}.$$

After one more derivative the ground state energy drops out of the problem (which is why I didn't define it in the problem statement):

$$\left(\frac{\partial^2 \ln Q}{\partial \beta^2} \right)_{N,V} = \frac{\epsilon^2 e^{-\beta \epsilon}(1 + e^{-\beta \epsilon}) - \epsilon^2 e^{-2\beta \epsilon}}{(1 + e^{-\beta \epsilon})^2} = \frac{\epsilon^2 e^{-\beta \epsilon}}{(1 + e^{-\beta \epsilon})^2}.$$

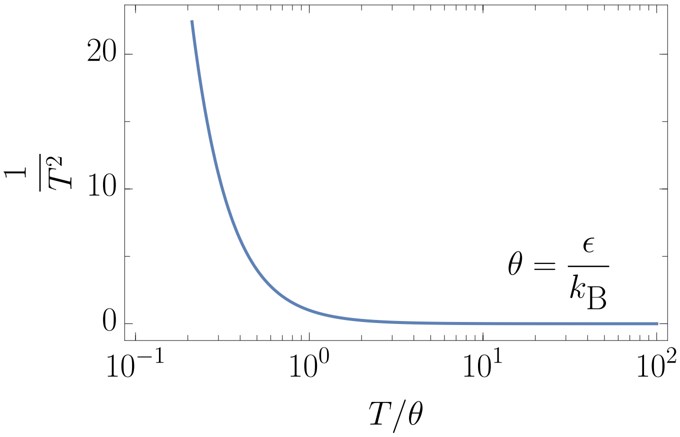
Inserting this derivative into our expression for C_V, we get the desired result:

$$C_V = \frac{\epsilon^2 e^{-\beta \epsilon}}{k_B T^2 (1 + e^{-\beta \epsilon})^2}.$$

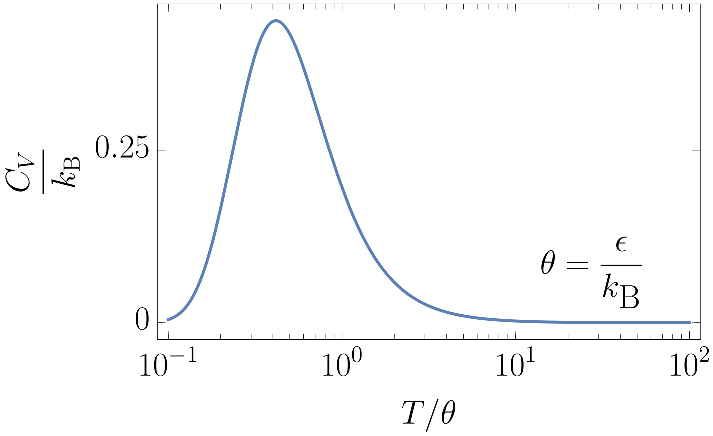
(iv) Combine results from parts (ii) and (iii) to express C_V in terms of k_B, T, and one or more cumulants of energy. **[3 pts.]**

$$C_V = \frac{\langle \delta E^2 \rangle}{k_\mathrm{B} T^2}.$$

(v) Use your answers to (ii) and (iv) to *roughly sketch* C_V. In making your sketch, you will probably find it useful to also think about the plot of $1/T^2$:



Do not worry about matching up the height of the y axis. Take it as given that I have put it on a reasonable scale so that your plot should fill up the space. **[4 pts.]**



Clearly, we see that the heat capacity is non-monotonic with T.

(vi) Hopefully by now your work has convinced me that the heat capacity can have non-monotonic features. But your work with the two-state model does not look exactly like the HD heat capacity. One difference is that the HD heat capacity jumps up multiple times; we can easily understand that difference as having to do with new classes of motion (rotations and vibrations) which are unlocked at higher temperatures. The bigger difference is that the high temperature limits do not seem compatible. The HD heat capacity per mole plateaus at 3.5 k_B while your two-state model heat capacity has a different limit. Explain the origin of the difference. **[5 pts.]**

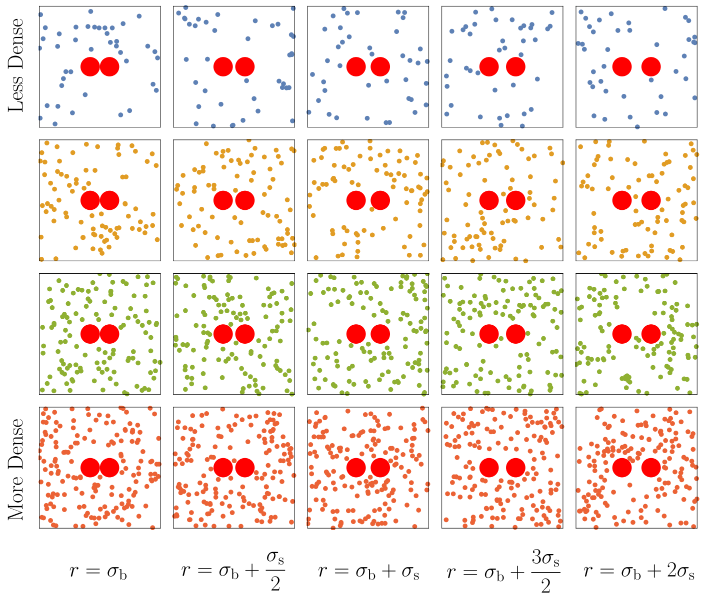
The two-state model has bounded energy levels, but the hd molecule does not. In the high-temperature limit, the hd molecule has, in principle, an infinite number of states that it can access. Recalling the definition of the heat capacity as the derivative of energy with respect to temperature, this implies that the energy can keep rising (at a constant rate) as the temperature rises. More importantly, the fluctuations can continue to increase without bound as t increases. By contrast, in the two-state model, the energy fluctuations cannot increase without bound but is capped by $\epsilon^2/4$. Hence the high-temperature limit of the two-state model decays to zero like $\epsilon^2/(k_\mathrm{B}t^2)$.

5. **Hard Spheres. [27 pts.]** This year's Nobel Prize in Chemistry was awarded for quantum size effects, including in colloidal nanoparticles. While much of the interest in those colloidal nanoparticles has focused on their optical properties, it is also been interesting to consider how multiple colloids interact with each other. As roughly spherical crystals, you might reasonably approximate a nanoparticle as if it is just a "hard sphere". Like billiard balls, such a model would express (a) that is is impossible for two nanoparticles to overlap each other and (b) that two nanoparticles separated by a gap do not exert forces on each other. Mathematically, one would express these facts with the pair potential:

$$V(r) = \begin{cases} \infty, & r < \sigma \\ 0, & r > \sigma, \end{cases}$$

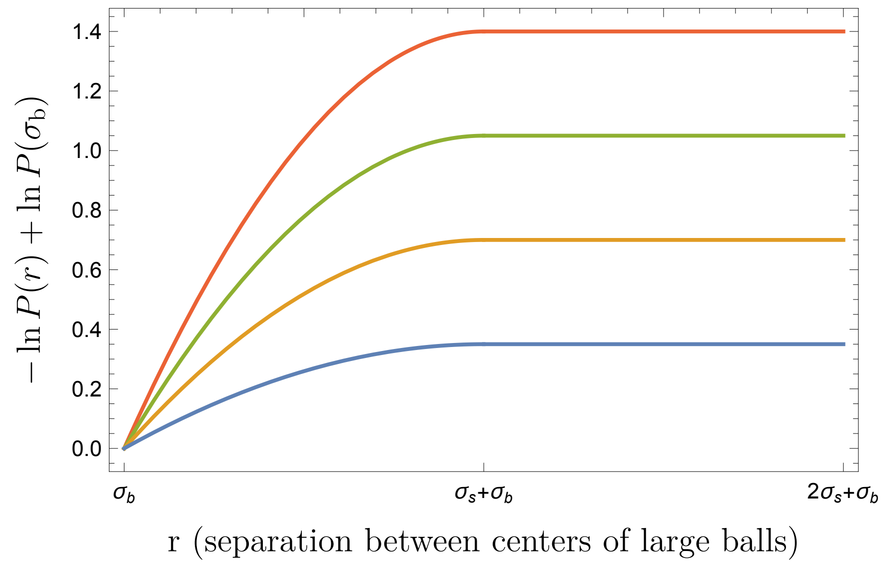
where r is the distance separating the center of the two spheres and σ is the diameter of the spheres. Models of hard spheres are most natural in two dimensions (hard disks) or three dimensions (hard spheres), though sometimes they're even considered in higher dimensions. In this problem, we will discuss an important statistical mechanical consideration for how two nanoparticles can interact in two-dimensional space.

We already said that when the two nanoparticles are separated by $r > \sigma$, they are not touching, so they don't really exert forces on each other. Suppose, however, that I put the hard disks in a solution of other smaller disks. To keep track of the difference, I will now call my comparatively big "nanoparticle" disk diameter σ_b and my comparatively small "solvent" disk diameter σ_s. The following picture gives you a sense of what configurations of the red nanoparticles look like in the solution of small solvent disks (colored by solvent density). Along the x axis I am adjusting the separation between the nanoparticles and along the y axis I am adjusting the density of the solvent particles.



Recall that we learned that it is much easier to work with models that have non-interacting particles, so let us treat the solvent as if each solvent disk is non-interacting with each other. In other words, multiple solvent disks can overlap each other (they are essentially an ideal gas). *However*, the solvent disks are assumed to still be hard with respect to the nanoparticles–small disks and large disks cannot overlap. The beauty of this simplification is that the impact of the solvent on the nanoparticles can be considered *one solvent disk at a time* without worrying about interactions between solvent disks.

(i) It is possible to exactly compute an NVE probability of each microscopic configuration, and to marginalize that distribution to get the probability $P(r)$ of finding the center of two nanoparticle disks separated by a distance r:



The four different lines on the plot are color coded to match the solvent densities of the configurations on the first page. According to the plot, does the solvent density affect the probability that the nanoparticles would be touching? Why or why not? **[5 pts.]**

The tricky part of this question was understanding what I was doing to normalize the probability distribution $P(r)$. For each density, the y axis was plotting that probability compared to the probability of touching disks: $P(\sigma_\mathrm{b}$. By construction, all four plots therefore pass through he same point at σ_b, but that does not mean that the probability of touching is the same for each density. A good way to approach this problem is to compare the probability at two different choices of r. For example, we could consider $P(\sigma_\mathrm{s} + \sigma_\mathrm{b})/P(\sigma_\mathrm{b})$ at each density. That consideration tells you that the height of the plot is measuring (on a log scale) how much less likely it is to find the separation $\sigma_\mathrm{s} + \sigma_\mathrm{b}$ than the separation σ_b. The higher the plot rises up, the less likely the larger separation is. Once you see that the relative probability between touching and not touching is strongly dependent on density, it follows that the probability of touching goes up as the solvent density goes up.

(ii) Suppose the nanoparticles were initially separated by a distance $\sigma_\mathrm{b} + \sigma_\mathrm{s}$. How much reversible work would I need to apply to separate them further? Does your answer depend on the solvent density? Justify both answers. **[5 pts.]**

The reversible work would be 0 and would not depend on the solvent density. Recall that the probability distribution for the spontaneous fluctuations (here $P(r)$) are related to the reversible work to adjust the single degree of freedom r by the reversible work theorem. This says that the difference in $-\ln P(r)$ from r_a to r_b is the reversible work, but we see that $-\ln P(r)$ is completely flat once we get to an r that is greater than or equal to $\sigma_\mathrm{b} + \sigma_\mathrm{s}$. That flatness, which was there at each of the densities, means there is zero reversible work for all densities.

(iii) Suppose the nanoparticles were initially touching each other and you wanted to pull them apart. Describe how one would use the plot to determine the equilibrium mean force you would need to apply to pull them apart. Without solvent, the force would have been zero. The nonzero force in the presence of solvent is an example of a "fluctuation-induced effective force" that emerges. The two nanoparticles do not *directly* exert force on each other, but due to interactions with the solvent they *effectively* do exert force. **[5 pts.]**

I wanted you to argue that the slope of the plots at $r = \sigma_\mathrm{b}$ would be the equilibrium mean force. You may have been a bit tripped up comparing to the reversible work theorem to figure out what to do about β. That is a tricky thing here (see the last part of the problem). For this part of the problem, if you described anything about getting equilibrium mean forces from that slope, I was happy.

(iv) Give an entropic argument to explain why the density of the solvent impacts the amount of reversible work it would take to separate the initially touching particles. [Hint: I am not merely asking you to use the plotted $P(r)$ here; I am asking of a physical argument you could have given even without seeing those explicit plots. The pictures of microscopic configurations may help some.] **[7 pts.]**

This is the hardest part of this problem. At zero density, the nanoparticles are just hard disks that do not exert forces on each other. Therefore, the reversible work to separate them would be zero (no forces). Once the small solvent disks are introduced, there is an entropic consideration that generates the effective force. This comes about because some configurations of the nanoparticles allow more configurations of the solvent than others. First, consider the two nanoparticles spread far apart. There is an empty annulus running from $r = \sigma_b$ to $r = \sigma_b + \sigma_s$ that wraps around the nanoparticle disk. Inside that annulus, no solvent disks can be centered because they would overlap with the nanoparticle disk. We might call it an exclusion annulus because solvent is excluded from being there. Each of the two nanoparticles has such an exclusion annulus, and they each limit how much volume is accessible to the "ideal gas" solvent disks. If, however, the two nanoparticles get close to one another, their two exclusion annului overlap so that solvent is excluded from a smaller total volume than if the nanoparticles had been far apart. That means the solvent has more arrangements (higher entropy) if the nanoparticles happen to be close together. With every additional solvent disk, that difference in the accessible solvent volume gets amplified, making the entropy difference between close nanoparticles macrostate and the far nanoparticles macrostate grow more dramatically different as the solvent density goes up. As the probability of the macrostates become more different, the amount of reversible work I must do to convert from one to the other also grows.

(v) Suppose a new Mirkin group member asks you how to make their nanoparticles stick together without altering the nanoparticles. Now that you've worked through this problem, you suggest introducing some polystyrene beads that are even smaller than the nanoparticle. Does this hard sphere model provide insight you can share with them about the temperature-dependence of the effective force between nanoparticles? If so, what is the insight? If not, why not? **[5 pts.]**

This problem is essentially a trick question. The hard sphere model is quite strange in that the energies are either zero or infinite. There is nothing in between. Consequently, if you ever want to talk about a Boltzmann distribution $e^{-\beta E}$, there can't be a difference between one value of

β and another. For all β, βE is still either 0 or ∞. One would say "hard spheres are athermal", so this model cannot really give me insight into the temperature dependence of the so-called "depletion attraction" between nanoparticles. I would have needed some model with a softer potential describing interactions between the particles (something that smoothly goes between 0 and ∞).