1. **Pulling on a Polymer.** Consider a model of a classical polymer in three dimensions, which consists of $N+1$ monomers connected together in a linear chain by N harmonic springs. We use $\boldsymbol{R}_i$ to denote the position of the i^{th} monomer. The energy of the polymer has kinetic and potential energy contributions, which depend on the positions and velocities of every monomer:

$$E(\boldsymbol{R}_0, \ldots, \boldsymbol{R}_N; \dot{\boldsymbol{R}}_0, \ldots, \dot{\boldsymbol{R}}_N) = \frac{k}{2} \sum_{i=1}^{N} |\boldsymbol{R}_i - \boldsymbol{R}_{i-1}|^2 + \frac{m}{2} \sum_{i=0}^{N} \left|\dot{\boldsymbol{R}}_i\right|^2,$$

where m is the mass of a monomer, $\dot{\boldsymbol{R}}_i$ is the velocity of monomer i, and the strength of the spring connecting polymers is k with units of energy per length squared. The polymer is immersed in a solution which has temperature T, so the energy of the polymer can fluctuate. The first part of the problem concerns this free polymer where "free" indicates that we are not applying extra forces to hold the endpoints of the polymer fixed in space.

Later in the problem we will imagine fixing monomer 0 at the origin by, for example, attaching it to a glass bead and holding that bead at the origin with a pipette. Monomer N will be fixed at $L\hat{x}$, which could be physically realized by attaching that monomer to a different glass bead and moving the bead with an optical trap. You will not need to know anything about how such an optical trap functions, only that it is possible to move the two endpoints of the polymer relative to each other so that they are separated by a distance L.

(i) Compute $\langle E \rangle$ as a function of k_{B}, T, and N. Your final answer should not depend on k or on m. (The result is a realization of the equipartition theorem.)

> Hint: To do this problem more efficiently, you may note that every bond between monomers is independent of the other bonds. The average energy in any one of these bonds can be expressed as a Gaussian integral,
>
> $$\left\langle E_{\text{single bond}} \right\rangle = \frac{k}{2} \left\langle R^2 \right\rangle = \frac{k}{2} \left(\frac{\beta k}{2\pi}\right)^{3/2} \int d\boldsymbol{R}\, R^2 e^{-\frac{\beta k R^2}{2}},$$
>
> where $\boldsymbol{R}$ is the vector (in three dimensions!) pointing from one monomer to the next and $\beta = 1/k_{\text{B}}T$. Similarly, the average kinetic energy of a single monomer is the same for all monomers and is given by
>
> $$\left\langle E_{\text{single monomer KE}} \right\rangle = \frac{m}{2} \left\langle \dot{R}^2 \right\rangle = \frac{m}{2} \left(\frac{\beta m}{2\pi}\right)^{3/2} \int d\dot{\boldsymbol{R}}\, \dot{\boldsymbol{R}}^2 e^{-\frac{\beta m \dot{R}^2}{2}},$$
>
> where $\dot{\boldsymbol{R}}$ is the velocity of the monomer.

(ii) Let us define $\boldsymbol{R}_{0\to N}$ to be the vector from monomer 0 to monomer N. Still talking about the free polymer, what is the ratio of probability densities $P(\boldsymbol{R}_{0\to N} = L\hat{x})/P(\boldsymbol{R}_{0\to N} = 0)$? In other words, how much less likely is it to find the stretched out free polymer than the compact one? Express your answer in terms of L, β, and N. Do the velocities matter? Why or why not? Hint: you are encouraged to make use of anything you learned in Problem Set 3 without having to rederive it here.

(iii) The polymer prefers to be folded up because there are fewer elongated configurations. Suppose the polymer starts out with both monomers 0 and N at the origin. What is the reversible work required for the optical trap to slowly move monomer N a distance L away, while keeping monomer 0 fixed at the origin?

2. **Reversible work simulations.**

   I have written a simulation of the polymer in Problem 1 being pulled apart at a finite rate (so not necessarily reversibly). You can run the simulation on this website:

   http://gingrich.chem.northwestern.edu/teaching/polymer/gaussianchain.html

   Each time you run the simulation you get a trajectory of the chain of harmonically bound nearest neighbors beads, with velocities of each bead assigned at the start of each trajectory from a Maxwell-Boltzmann distribution. Each bead of the polymer experiences forces from the neighboring beads, trying to pull the neighbors closer together, but it also feels random "kicks" from the environment. These kicks are meant to mimic the forces from molecules in a fixed-temperature solvent colliding with the polymer. In the simulation, the solvent is not explicitly simulated, rather, at each step of the polymer's dynamics, a random Gaussian number is drawn by the computer to represent the extra force a monomer feels from the effect of the solvent. Finally, the endpoint of the polymer experiences an external force that we apply to force the polymer to spread out at a fixed stretching rate. This applied force is measured and plotted in the top right of the screen. When the applied force is integrated, we compute the total work exerted by the external force, and this accumulated work is plotted in the bottom right of the screen.

   If you slow down the pulling rate, you will see that the total work to stretch out the chain matches the reversible work.

   (i) First consider the case of an infinitesimally slow pulling rate. Compute the average force that must be applied on the end monomer as a function of k, L, and N. This should follow from your answer to Problem 1, Part (iii).

   (ii) Use the provided simulation, starting with $N = 20$ beads and a pulling rate of 0.1 (in reduced units where $\beta = 1, l = 1$, and bead mass $m = 1$). Then try a pulling rate of 0.01 and a pulling rate of 0.001. Observe that for very slow pulling the total work is very close to the reversible work. Notice also that the mean force fluctuates around the correct average value that you found in (i). But the fluctuations in the measured force are quite large, even at slow pulling rates. Explain why the force fluctuations do not decrease for slow pulling rates even though the work fluctuations do.

3. **A perpetual motion machine?!** Your friend at the University of Chicago has come up with a brilliant idea. When he played with the simulation, he set $N = 15$ and pulled with a rate of 0.1 (in the reduced units). He observed that sometimes he gets negative work by stretching the polymer out. In other words, he expected to have to do work to pull the polymer apart, but sometimes the polymer did work on him. Your friend is super excited because he thinks he can build some sort of DNA/laser tweezer machine that stretches and unstretches a strand of DNA and gets work out in the process. Use the simulation posted at:

   http://gingrich.chem.northwestern.edu/teaching/polymer/cyclicworkdistribution.html

   to help you evaluate his plan.

   (i) During a single cycle, will the UChicago scientist ever extract work from the system (the measured work is negative)?

   (ii) Will the average extracted work ever be negative? To answer, compute the average work $\langle W \rangle$ for pulling rates of $0.05, 0.1,$, and 0.2. You will want to check "Repeat Pulling to Collect Work Statistics"

then click restart to automatically generate statistics of the pulling experiments. Make sure you run the program long enough for the values of $\langle W \rangle$ to converge. What do you notice about how the distribution of measured work values depends on the pulling rate? Will there be a pulling rate so that your friend's machine will reliably extract work and make him billions of dollars?

(iii) Based upon the time-reversibility of dynamics (i.e., Newton's laws look the same forward as they do backwards), it can be shown that the probability of observing a value of work, W, over one of the cycles is related to the probability of measuring $-W$ by

$$\frac{P(W)}{P(-W)} = e^{\beta W}. \tag{1}$$

(This formula is a special case of the Crooks fluctuation theorem, which we will discuss further in the next problem.) At a single pulling rate, run the simulation long enough to generate a smooth work probability distribution. Once the distribution has converged use the raw data provided below the plot to confirm that the probability of positive and negative values of work are related as predicted by Equation (1). (Agreement will be good but not perfect only because of finite statistics and because histogram bin sizes analyzed in the applet are fairly wide.) In the reduced units of the simulations, $\beta = 1$.

(iv) The average work from Part (ii) can be considered to be an average over the work distribution function, $P(W)$.

$$\langle W \rangle = \int_{-\infty}^{\infty} dW \, W P(W)$$

Show that Equation (1) implies that your answer to Part (ii) must be positive.

(v) Beginning with the fact that $P(W)$ is normalized, show that $1 = \left\langle e^{-\beta W} \right\rangle$. The average in this equation is over all possible values of work, which means averaging over all of the possible non-equilibrium pulling experiments.

(vi) This average of the exponential is computed in the simulation and reported on the screen. Note that it converges to one, but the convergence is slow. Comment on why convergence is so slow. [Hint: Think about how probable it is to observe measurements that contribute the most to the average.]

4. **Equilibrium free energies from nonequilibrium measurements.** In the previous problem, you considered a cyclic process in which the polymer endpoints returned to their original positions at the end of each run. If the reversible work between the starting and ending configurations is not zero, the general form of the Crooks fluctuation theorem applies:

$$P_{\rm F}(W) = P_{\rm R}(-W) e^{\beta W_{\rm d}} \, . \tag{2}$$

Here, $W_{\rm d} = W - W_{\rm rev}$ is the extra work we do on top of the reversible work. $P_{\rm F}(W)$ refers to the probability of measuring a value W when pulling the polymer. $P_{\rm R}(-W)$ is the probability of measuring a value of $-W$ when *pushing* the endpoints of the polymer back to where they started. (F stands for forward process, R for reversed.) In experiments with DNA folding it is important to know the reversible work for unfolding a piece of DNA, but we cannot pull the DNA slowly enough to unfold the polymer reversibly. Averaging the work performed during a laser tweezing experiment is a bad estimate for $W_{\rm rev}$ because $\langle W \rangle \neq W_{\rm rev}$. Notice that the Crooks fluctuation theorem establishes that the difference in statistics between forward and backward processes – the nature of time's arrow – is determined entirely by value of the dissipated work, $W_{\rm d}$.

Repeat the logic of Problem 3(v) to write down an expression for $W_{\rm rev}$ in terms of the average of some quantity you could measure in the pulling experiments. [Hint: $P_{\rm R}(W)$ is a normalized probability distribution.]

You can check that this gives a good estimate for W_{rev} with using the simulation posted at:

http://gingrich.chem.northwestern.edu/teaching/polymer/workdistribution.html.

With the Crooks fluctuation theorem, therefore, you have a way to extract measurements that tell us about equilibrium systems (W_{rev}) from experiments that are not even close to equilibrium!