Lecture 19: Polymer Models: Persistence and Self-Avoidance

Scribe: Allison Ferguson (and Martin Z. Bazant)

Martin Fisher School of Physics, Brandeis University

April 14, 2005

Polymers are high molecular weight molecules formed by combining a large number of smaller molecules (monomers) in a regular pattern. Polymers in solution (i.e. uncrystallized) can be characterized as long flexible chains, and thus may be well described by random walk models of various complexity.

1 Simple Random Walk (IID Steps)

Consider a Rayleigh random walk (i.e. a fixed step size a with a random angle θ). Recall (from Lecture 1) that the probability distribution function (PDF) of finding the walker at position $\vec{R}$ after N steps is given by

$$P_N(\vec{R}) \sim \frac{e^{\frac{-3R^2}{2a^2N}}}{(2\pi a^2 N/d)^{\frac{d}{2}}} \tag{1}$$

Define $R_N = |\vec{R}|$ as the end-to-end distance of the polymer. Then from previous results we know $\langle R_N^2 \rangle = Na^2$ and $\bar{R_N} = \sqrt{\langle R_N^2 \rangle} = a\sqrt{N}$. We can define an additional quantity, the radius of gyration G_N as the radius from the centre of mass (assuming the polymer has an approximately spherical shape[1]). Hughes [1] gives an expression for this quantity in terms of $\langle R_N^2 \rangle$ as follows:

$$\langle G_N^2 \rangle = \frac{1}{6}\left(1 - \frac{1}{N^2}\right)\langle R_N^2 \rangle \tag{2}$$

As an example of the type of prediction which can be made using this type of simple model, consider the following experiment: Pull on both ends of a polymer and measure the force f required to stretch it out[2]. Note that the force will be given by $f = -(\partial F/\partial R)$ where $F = U - TS$ is the Helmholtz free energy (T is the temperature, U is the internal energy, and S is the entropy). In the simplest random-walk model of a polymer chain, we neglect forces between monomers (and the solvent), aside from the constraint of connecting the monomers in a chain, so there is no internal energy, $U = 0$. The free energy is thus fully determined by the entropy, S, which essentially

[1] As described in Rudnick and Gaspari [2], in general, the typical shape of a random walk is better represented by an ellipsoid. This is a consequence of the non-uniform exploration of space by a single walker, related to the "arc-sine law" from problem set 1.

[2] This experiment has actually been done on DNA, by attaching glass beads to the each end of the polymer and using optical tweezers to exert force on the beads, thus stretching the DNA[3].

measures the number of available states, each assumed to be equally likely in thermal equilibrium, subject to the constraint of fixed end-to-end distance R. Using the definition, $S = k_b \log P_N$, the entropy can be calculated for the Rayleigh random walk (taking only dominant terms for $N \gg 1$) as

$$S \simeq -k_b \frac{3R^2}{2a^2 N} \tag{3}$$

Then the free energy is given by

$$F = -TS \simeq \frac{3}{2} k_b T \frac{R^2}{a^2 N} \tag{4}$$

Letting $k_s \simeq 3k_b T/Na^2$ we can see that this has the familiar form of the simple harmonic oscillator potential $F = (1/2)k_s R^2$. Thus the force required to stretch a polymer is $f = -k_s R$ and the polymer is modelled by a simple linear (Hooke) spring. Note that the "spring constant" k_s predicted by this model is proportional to the temperature T: At high temperature, the polymer has a strong tendency to remain in a densely coiled state with high degeneracy, while at low temperature, the polymer is easily stretched out toward the non-degenerate linear state.

Of course, something is amiss in the model as $R \to \infty$ since the constraint, $R \leq Na$, is missing (i.e. the polymer cannot be longer than $R = Na$, corresponding to a perfect linear chain). However, this should come as no surprise since our expression for the entropy is only valid in the "central region", $R = O(a\sqrt{N})$. In principle, one should use a globally valid approximation from saddle point asymptotics (lecture 7) to better describe the entropty of highly stretched configurations. The relevant case of the Rayleigh walk is discussed by Hughes [1].

2 Persistent Random Walk

In reality, the assumption of uncorrelated steps does not accurately represent many common polymers. For example, in a protein with a carbon backbone, nearest neighbouring bonds have a preferred angle due to the tetrahedral structure formed by the carbon atoms[3] Thus a better model may involve *correlated* steps.

We can incorporate this constraint in the random walk, $\vec{X}_n = \sum_{n=1}^{N} \vec{x_n}$, by assuming a direct correlation coefficient ρ between sucessive displacements,

$$\langle \vec{x_n} \cdot \vec{x_{n+1}} \rangle = \rho \tag{5}$$

where $-1 < \rho < 1$. Here, $\rho = 0$ corresponds to the uncorrelated case, and $\rho = 1/3$ would represent the case of the protein backbone described above. When $\rho > 0$, the walker has a tendency to keep moving in the same direction, which is sometimes referred to as "persistence". The persistent random walk was introduced by G. I. Taylor in 1921 as a model for diffusion by chaotic advection in a turbulent fluid flow, taking into account the non-diffusive linear motion of a tracer particle in the flow at small time scales.

As before, let us assume $\langle \vec{x_n} \rangle = 0$ and $\langle \vec{x_n}^2 \rangle = \sigma^2$. In the next lecture[4], we will show that the persistent walk still exhibits "normal diffusion" (variance asymptotically linear in the number of steps), but with a modified diffusion coefficient:

$$\langle \vec{X_n}^2 \rangle \sim N\sigma^2 \left(\frac{1+\rho}{1-\rho} \right) = N\sigma_{eff}^2 \tag{6}$$

[3]This is a consequence of four sp^3 hybrid covalent bonds per carbon atom: two which form connections with neighboring carbon atoms in the backbone and two which form bonds with sidebranches and various attached molecules. The tetrahedral angle between adjacent carbon atoms is $\cos^{-1}(-1/3) \approx 109$ degrees.

[4]See also 2003 18.366 notes, lecture 9.

Note that for $0 < \rho < 1$ then $\sigma_{eff} > \sigma$ (σ_{eff} diverges as $\rho \to 1$). Thus the effect of persistence ($\rho < 0$) is to increase σ_{eff}. For our protein example, $\rho = 1/3$ and thus

$$\sigma_{eff}^2 = a^2 \left(\frac{1+\rho}{1-\rho} \right) = 2a^2 \tag{7}$$

Since $\bar{R_N} = \sqrt{2}a\sqrt{N}$, we recover the same scaling as for IID steps, but the step size is effectively larger by a factor of $\sqrt{2}$.

3 Self Avoiding Walk

For real polymers, the "walk path" has finite thickness and cannot self-intersect. For a random walk on a lattice, this would mean that the walk can visit a given lattice site only once, but more generally, we could consider an off-lattice walk, where a sphere of diameter a is excluded for subsequent steps. Self avoidance is a rather non-trivial example of long-range correlations, where the next step of the walk dependece on the entire previous history of the walk. As such, analytical progress is difficult (especially in the most relevant case of $d = 3$ dimensions), and most of what we know about self-avoiding walks comes from computer simulations. See Hughes for a detailed discussion [1].

3.1 Anomalous scaling

One effect of self-avoidance is to introduce a different scaling exponent for the average end-to-end distance of the polymer $\bar{R_N} = aN^\nu$, where $\nu > 1/2$ (which would correspond to "superdiffusion"). The scaling exponent ν is clearly dependent on the dimension of space: For $d = 1$, self-avoidance causes the walk to always move in one direction, yielding $\nu = 1$, but as $d \to \infty$ we expect to recover $\nu = 1/2$ since in higher spatial dimensions the effects of self-avoidance become less important, as the return probability goes to zero. (Recall Pólya's theorem.) Indeed, analytical and numerical calculations made using lattice models have yielded the following results for ν:

$$\nu = \begin{cases} 1 & d = 1 \\ 3/4 & d = 2 \\ 0.586 & d = 3 \\ 1/2 & d \geq 4 \end{cases}$$

For $d = 2$, the result is analytical and considered to be "exact", but for $d = 3$ it comes only from simulations. The result $\nu = 1/2$ is exact for "$d = \infty$" (when the return probability vanishes), but it is believed to also hold down to some finite dimension, $d \geq d_c$, where d_c is the *upper critical dimension*, believed to be four for this problem[5].

3.2 Flory's Theory

One of the earliest theoretical descriptions of self-avoidance was constructed by Nobel Prize-winning chemist Paul Flory [4] in the 1940s. This remarkably effective mean-field model reduces the complicated many-body problem of the conformational sum over all the monomer orientations with interactions, to a simple approximate estimate of the net interaction energy as a function of R by

[5] For more information on how upper critical dimensions are determined and on the construction of mean-field theories in general please see Ref. [5]

introducing an additional parameter E (E is essentially an "energy penalty" for revisiting a given lattice site). Then this simplified expression for the interaction energy is added to the stretching energy we derived at the beginning of the lecture by considering only IID steps. Flory's expression for the total free energy of a polymer with self-avoidance as a function of R is constructed as follows: Imagine N monomers isotropically distributed in a volume R^d. Then by considering a mean field description of reaction rates [6] we can write down the interaction energy U as

$$U \propto \frac{EN^2}{R^d} \tag{8}$$

Therefore the total free energy (using the previously determined result for TS) is

$$F = A\frac{N^2}{R^d} - B\frac{R^2}{N} \tag{9}$$

where A and B are constants independant of R and N. By minimizing Eq. 9 with respect to R, the scaling behaviour of R with N can be determined as:

$$R \propto N^{\frac{3}{2+d}} \tag{10}$$

The accuracy of Flory's prediction,

$$\nu = \frac{3}{2+d}, \quad d \leq 4 \tag{11}$$

is truly remarkable. It reproduces the *exact* results for $d = 1, 2, 4$, and is within one percent of simulation estimates for $d = 3$. Note that the formula agrees with the limit of no self-avoidance, $\nu = 1/2$, at $d = d_c = 4$, which is thus predicted to be the upper critical dimension. On the other hand, the formula makes no sense for $d > 4$, since the effect of self avoidance can only increase the scaling exponent from the independent case, $\nu > 1/2$.

References

[1] B. Hughes, *Random Walks and Random Environments Volume 1* (Oxford, 1996)

[2] J. Rudnick and G. Gaspari *Elements of the Random Walk* (Cambridge, 2004)

[3] C. Bustamante et. al., Science **265**, 1599 (1994)

[4] P. J. Flory *Principles of Polymer Chemistry* (Cornell University Press, 1953)

[5] P. M. Chaikin and T. C. Lubensky, *Principles of condensed matter physics* (Cambridge, 1995)

[6] If a chemical reaction is described by the rate equation $mA + nB \rightarrow C$, then the "mean-field" reaction rate (assuming no positional correlations between reactants) is given by $\rho_A^m \rho_B^n$.