

## 2.2 Fluctuating Polymers

The basic molecules of life (DNA, RNA, proteins, ...) are *hetero-polymers*, formed by the covalent bonding of a sequence of elementary units (nucleic acids, amino-acids) in long chains. A *homo-polymer*, as in many synthetic organic molecules, is constructed by joining  $N \gg 1$  copies of the same monomer. A simple example is *polyethylene*,

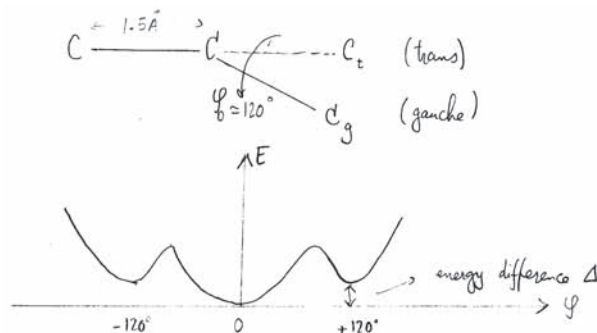


The *degree of polymerization*, i.e. the the number of repeated units, can be quite large, ranging from a few hundred for proteins,  $10^4 - 10^5$  for polyethelene, to as big as  $10^9$  for some DNA. Typically the covalent bonds holding the polymer together are strong and cannot be broken at room temperature. There can, however, be flexibility in aligning/bending successive monomers, resulting in a large number of configurations degrees of freedom for polymers, indicating that a statistical description of the problem is fruitful. Such a statistical perspective is useful for describing general properties common to both synthetic and natural polymers. For example, at very high (not necessarily physiological) temperatures all polymers will be in a swollen (denatured) state to take maximum advantage of entropy. The heterogeneity of the sequence is irrelevant in such a phase, and we shall thus initially focus on the fluctuations of a homo-polymer.

### 2.2.1 Rotational isomers

Successive carbon-carbon bonds in the chain can be in different relative orientations, called *rotational isomers*. In polyethelene, the low energy *trans* conformation leads to a parallel alignment of bonds. There are two higher energy *gauche* states which form an angle of  $2\pi/3$  between successive bonds. Assuming an energy difference  $\Delta$  between the trans, and the gauche states, at a temperature  $T$ , the probabilities of these outcomes satisfy the Boltzmann weight

$$\frac{\text{prob.}(g)}{\text{prob.}(t)} = 2e^{-\beta\Delta}, \quad \text{with} \quad \beta = \frac{1}{k_B T}.$$



For  $(CH_2)^N$ ,  $\Delta$  is quite small (roughly 500 cal/mole or  $1/3k_B T$ ), and the polymer is very flexible at room temperature. For other polymers  $\Delta > k_B T$ , and gauche states with probability

$$p(g_+) = p(g_-) = \frac{e^{-\beta\Delta}}{1 + 2e^{-\beta\Delta}},$$

are relatively rare. A typical configuration then consists of long straight segments with few bends. The probability of a straight segment of  $n$  monomers is  $p_t^n (1 - p_t)$ , where  $p_t = 1 - 2p(g)$  is the probability of a trans bond. The average length of straight segments is thus  $\langle n \rangle a$ , where  $a$  is the bond length (monomer size), and

$$\langle n \rangle = -\frac{1}{\ln p_t} = (\ln [1 + 2e^{-\beta\Delta}])^{-1}. \quad (2.28)$$

The typical size of these linear segments is proportional to the *persistence* of the polymer. The *persistence length* characterizes the decay of orientational correlations along the chain. In the above simplified model, let us denote the orientation of the bonds by the set of vectors  $\{\vec{t}_1, \vec{t}_2, \dots, \vec{t}_N\}$ , with  $\vec{t}_j \cdot \vec{t}_j = a^2$ . Assuming that the (two) gauche states produce a relative bond angle  $\phi$ , we have

$$\langle \vec{t}_1 \cdot \vec{t}_2 \rangle = a^2 \frac{1 + 2 \cos \phi e^{-\beta\Delta}}{1 + 2e^{-\beta\Delta}} \approx a^2 \exp[-2(1 - \cos \phi)e^{-\beta\Delta}],$$

where the last expression is valid in the limit of  $\beta\Delta \gg 1$  where a gauche state is very unlikely. In the same approximation, the correlation between bonds that are further apart is given by

$$\langle \vec{t}_1 \cdot \vec{t}_{n+1} \rangle \approx a^2 \frac{1 + 2n \cos \phi e^{-\beta\Delta} + \mathcal{O}(e^{-2\beta\Delta})}{1 + 2ne^{-\beta\Delta} + \mathcal{O}(e^{-2\beta\Delta})} \approx a^2 \exp[-2n(1 - \cos \phi)e^{-\beta\Delta}],$$

where we have included only configurations with one gauche bond. The orientation correlations decay exponentially as  $e^{-\ell/\xi_p}$ , where  $\ell = na$  is the counter-length along the polymer, and the persistence length  $\xi_p$  is given by

$$\xi_p \approx \frac{ae^{\beta\Delta}}{2(1 - \cos \phi)}. \quad (2.29)$$

### 2.2.2 Worm-like chain

For a rigid polymer such as double-stranded DNA a kink causing a finite rotational angle is energetically costly and does not occur. The loss of angular correlations at long distances then occurs from the accumulation of small changes from one monomer to the next. If we indicate as before a polymer configuration by the set of vectors  $\{\vec{t}_1, \vec{t}_2, \dots, \vec{t}_N\}$ , we can approximate the energy of a nearly straight configuration by

$$\mathcal{H} = -J \sum_{i=1}^{N-1} \vec{t}_i \cdot \vec{t}_{i+1}, \quad (2.30)$$



Since in a typical configuration  $\vec{t}$  changes slowly, it is useful to go over to a *continuum limit* in which the discrete monomer index  $i$  is replaced by the continuous *arc-length*  $s \in [0, L = Na]$ . Using  $(\vec{t}_i - \vec{t}_{i+1})^2 = 2 - 2\vec{t}_i \cdot \vec{t}_{i+1}$ , we then can write

$$\mathcal{H} \approx -JN + \frac{\kappa}{2} \int_0^L ds \left( \frac{d\vec{t}}{ds} \right)^2, \quad (2.31)$$

where  $\kappa = J/a$  is the coefficient of *bending rigidity*. (Note that  $|d\vec{t}/ds| = 1/R(s)$ , where  $R(s)$  is the local radius of curvature.)

Ignoring the initial energy of the “ground state” configuration, it is common to write the energy in dimensionless form as

$$\beta\mathcal{H} = -\frac{\xi_p}{2} \int_0^L ds \left( \frac{d\vec{t}}{ds} \right)^2, \quad (2.32)$$

with  $\beta\kappa = \xi_p$ . We have anticipated that the bending rigidity is related to the persistence length. In fact, it can be shown (e.g. by using transfer matrices) that for the discrete model of Eq. (2.30)

$$\langle \vec{t}_m \cdot \vec{t}_n \rangle \approx \left( \coth(\beta J) - \frac{1}{\beta J} \right)^{|m-n|} \quad \text{for } |m-n| \gg 1, \quad (2.33)$$

and thus in the continuum limit

$$\langle \vec{t}(s) \cdot \vec{t}(s+\ell) \rangle \approx e^{-\ell/\xi_p}, \quad (2.34)$$

with  $\xi_p \approx \beta\kappa$  for  $\beta J \gg 1$ . This, so called *worm-like chain* model is frequently invoked as a description of double-stranded DNA, where  $\xi_p$  is in the range of 50-100nm.

### 2.2.3 Entropic elasticity

The flexibility of long polymers arises from the statistical fluctuations of segments larger than the persistence length. The important parameter that governs the number of configurations is thus not the degree of polymerization  $N$ , but the number of unconstrained degrees of freedom, or the *Kuhn* length  $N_K \approx Na/(2\xi_p)$ . To see this explicitly, let us consider the *end to end separation* of the polymer, given by

$$\vec{R} = \vec{t}_1 + \vec{t}_2 + \cdots + \vec{t}_N = \sum_{i=1}^N \vec{t}_i.$$

Because of rotational symmetry (there is no cost for rotating the entire polymer),  $\langle \vec{R} \rangle = 0$ , and its variance is given by

$$\langle R^2 \rangle = \sum_{i,j=1}^N \langle \vec{t}_i \cdot \vec{t}_j \rangle = Na^2 + 2 \sum_{i<j} \langle \vec{t}_i \cdot \vec{t}_j \rangle. \quad (2.35)$$

We shall assume that the orientational correlations decay as a simple exponential (this is only asymptotically correct), i.e.

$$\langle \vec{t}_i \cdot \vec{t}_j \rangle = a^2 e^{-a|i-j|/\xi_p}. \quad (2.36)$$

As the correlation function is the same for every pair of points separated by a distance  $k$ , and as there are  $(N - k)$  such pairs along the chain

$$\langle R^2 \rangle = Na^2 + 2a^2 \sum_{k=1}^N (N - k) e^{-ak/\xi_p}. \quad (2.37)$$

The geometric series are easily performed, and for  $N \gg 1$  (where only the term proportional to  $N$  is significant), we obtain

$$\langle R^2 \rangle = Na^2 \left( 1 + \frac{2e^{-a/\xi_p}}{1 - e^{-a/\xi_p}} \right) = Na^2 \coth \left( \frac{a}{2\xi_p} \right) \quad (2.38)$$

$$\approx 2Na\xi_p = (2\xi_p)^2 \left( \frac{Na}{2\xi_p} \right). \quad (2.39)$$

The approximations in the second line rely on  $\xi_p \gg a$ . The very last expression indicates that the behavior of the variance is the same as that of  $N_K \equiv (Na/2\xi_p)$  independent rods of length  $2\xi_p$ , i.e. the same variance is obtained for a collection of  $N_K$  *freely-jointed rods*, each of length  $2\xi_p$ . Indeed the correlations between these Kuhn segments is sufficiently small that in the limit of  $N_K \gg 1$ , we expect the *Central Limit Theorem* to hold, leading to the probability distribution function

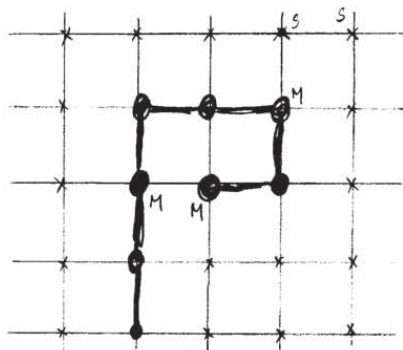
$$p(\vec{R}) = \exp \left[ -\frac{3R^2}{2\langle R^2 \rangle} \right] \left( \frac{2\pi\langle R^2 \rangle}{3} \right)^{3/2} = \exp \left[ -\frac{3R^2}{4Na\xi_p} \right] \frac{1}{(4\pi Na\xi_p/3)^{3/2}}. \quad (2.40)$$

The final probability distribution is identical to the Boltzmann weight of a Hookian spring of strength  $J$  connecting the end points of the polymer, and the result of entropic fluctuations can be interpreted as conferring an elastic bond between the ends of the polymer with spring coefficient

$$J_{polymer} = \frac{3k_B T}{\langle R^2 \rangle} = \frac{3k_B T}{2Na\xi_p}. \quad (2.41)$$

## 2.3 Interacting Polymers

The polymeric properties discussed so far arose from the flexibility of the covalent bonds that join adjacent monomers. There are also interactions between any other pairs (triplets, etc.) of monomers which depend on their spatial vicinity and that favor certain spatial configurations. Indeed, it is such interactions, typically due to hydrogen bonds, that enable proteins to fold and assume specific shapes. There are competing effects due to thermal fluctuations and competition with solvent interactions.



Some of the ingredients of polymer interactions in a solvent is present in the very simple model of chain configurations on a (say square) lattice. The set of *random walks* on the square lattice that do not step back to the previous site grows with the number of steps  $N$  as  $3^N$ . One simple consequence of interactions is that it is physically impossible to visit a previously occupied site. The constraint of *excluded volume* prunes the random walks to a smaller subset of so-called *self-avoiding walks*. The number of self-avoiding walks also grows exponentially as  $g^N$  with  $g < 3$  ( $g \approx 2.64$ ).

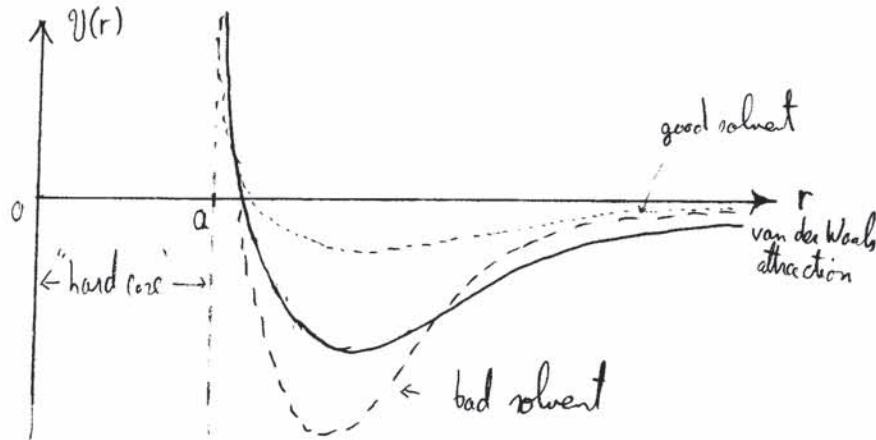
A simple way to incorporate interactions on a lattice is to count the number of (non-polymeric) nearest-neighbor pairs, and assign energy

$$E = \epsilon_{mm}N_{mm} + \epsilon_{ms}N_{ms} + \epsilon_{ss}N_{ss},$$

where  $mm$ ,  $ms$ , and  $ss$  stand for monomer-monomer, monomer-solvent, and solvent-solvent pairs respectively, with  $N_{pair}$  and  $\epsilon_{pair}$  indicating the corresponding number and bond-energies. As two initially separate monomers are brought into contact, two  $ms$  bonds are replaced by one  $mm$  bond and one  $ss$  bond, leading to a change in energy of  $\delta\epsilon = \epsilon_{mm} + \epsilon_{ss} - 2\epsilon_{ms}$ . The preference of the monomers to aggregate in solvent is thus captured by the dimensionless “*Flory-Huggins*” parameter

$$\chi \equiv -\frac{\beta}{2}\delta\epsilon = \beta \left( \epsilon_{ms} - \frac{\epsilon_{mm} + \epsilon_{ss}}{2} \right). \quad (2.42)$$

A negative  $\chi$  leads to separation of monomers, while a positive  $\chi$  encourages their aggregation.



In a more realistic model, the interactions between molecules vary continuously as a function of their relative separation and orientation in space. The dependence on orientation is particularly relevant to hydrogen bonding, while the van der Waals attraction mainly depends on the separation. Just as in Eq. (2.8), an effective potential  $\mathcal{V}(r)$  between monomers is in principle obtained by integrating over all positions (and orientations) of the solvent particles. In the usual case where the monomers are larger than the solvent molecules, the effective potential is attractive at large distances and has a hard repulsive core at short distances. For a good solvent the potential is weak, while a strong attractive potential favors aggregation of monomers in a bad solvent. The quality of solvent also changes as a function of temperature due to entropic contributions of its constituents. The larger entropy of solvent molecules typically improves the quality of a solvent at higher temperatures.

### 2.3.1 Mean-field estimate of the partition function

To calculate the properties of a polymer in solvent— e.g. to determine if it is in its native form or a denatured state at some temperature— we need to compute the free energy of the molecule and its environment. Computation of the partition function is a hard task, even for the highly simplified models we have introduced so far. We shall instead rely upon an approximate expression obtained in a mean-field/variational treatment. Let us assume that the most likely configurations of an interacting homopolymer have a typical size  $R$ . The partition of  $N$  monomers confined to a sphere of radius  $R$  is then estimated as

$$\begin{aligned}
 Z(N, R) &\approx g^N \times \frac{\exp\left(-\frac{3R^2}{4Na\xi_p}\right)}{(4\pi Na\xi_p/3)^{3/2}} \\
 &\times \left\{ 1 \cdot \left[1 - \left(\frac{a}{R}\right)^3\right] \cdot \left[1 - 2\left(\frac{a}{R}\right)^3\right] \cdots \left[1 - (N-1)\left(\frac{a}{R}\right)^3\right] \right\} \times e^{-\beta E_{att}}.
 \end{aligned}
 \tag{2.43}$$

The first line in Eq. (2.44) pertains to the entropy of the polymer, the first term encodes the exponential growth in the number of configurations of an unconstrained polymer— the

precise value of  $g$  is in fact irrelevant to the considerations that follow. The second term approximates the reduction in the number of configurations when the polymer is constrained to a size  $R$ . The effect of this reduction is included as a Hookian spring, motivated by the result in Eq. (2.40) for the end-to-end probability of a non-interacting polymer.

The second line in Eq. (2.44) approximates the effect of interactions and is broken into two parts. The first part encodes the reduction in phase space due to excluded volume constraints: the first monomer is unconstrained, the volume available to the second is reduced by the fraction  $(a/R)^3$  due to the volume excluded by the first, and so on. The reductions due to the excluded volume make a contribution to the free energy proportional to

$$\begin{aligned} \delta \ln Z_{EV} &= \sum_{i=1}^{N-1} \ln \left[ 1 - i \left( \frac{a}{R} \right)^3 \right] \approx - \left( \frac{a}{R} \right)^3 \sum i - \frac{1}{2} \left( \frac{a}{R} \right)^6 \sum i^2 + \dots \\ &\approx - \frac{N^2}{2} \left( \frac{a}{R} \right)^3 - \frac{N^3}{6} \left( \frac{a}{R} \right)^6 - \dots \end{aligned} \quad (2.44)$$

The attractive part of the interaction, for homopolymers, is given by

$$E_{att.} = \frac{1}{2} \sum_{i \neq j} \mathcal{V}(\vec{r}_i - \vec{r}_j) = \frac{1}{2} \int d^3 \vec{r} d^3 \vec{r}' n(\vec{r}) n(\vec{r}') \mathcal{V}(\vec{r} - \vec{r}'). \quad (2.45)$$

Assuming a uniform mean-density,  $n(\vec{r}) = n = N/V = N/(4\pi R^3/3)$ , leads to

$$E_{att.} = \frac{n^2}{2} V \int_a d^3 \vec{r} \mathcal{V}(\vec{r}), \quad (2.46)$$

where we have integrated over the center of mass of the pair to get the volume  $V$ , and ignored any contributions from the surface. In the spirit of Flory-Huggins, we introduce a dimensionless parameter  $\chi$ , via

$$\int_a d^3 \vec{r} \mathcal{V}(\vec{r}) \equiv \left( -\frac{4\pi}{3} a^3 \right) k_B T (2\chi), \quad (2.47)$$

to capture of the net effect of attractions, and such that

$$-\beta E_{att.} = N^2 \left( \frac{a}{R} \right)^3 \chi. \quad (2.48)$$

The resulting free energy, with  $R$  as a variational parameter,

$$\begin{aligned} \ln Z(N, R) &= N \ln g - \frac{3}{2} \ln \left[ \frac{4\pi N a \xi_p}{3} \right] - \frac{3R^2}{4N a \xi_p} \\ &\quad - \frac{N^2}{2} \left( \frac{a}{R} \right)^3 - \frac{N^3}{6} \left( \frac{a}{R} \right)^6 - \dots + \chi N^2 \left( \frac{a}{R} \right)^3. \end{aligned} \quad (2.49)$$

will next be used to explore the phases of the interacting homopolymer.

MIT OpenCourseWare  
<http://ocw.mit.edu>

8.592J / HST.452J Statistical Physics in Biology  
Spring 2011

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.