

# The conformations of linear polymers

Marc R. Roussel  
Department of Chemistry and Biochemistry  
University of Lethbridge

February 19, 2009

Polymer science is a rich source of problems in applied statistics and statistical mechanics. In these notes, we will look at the conformations of polymers from this perspective.

## 1 Mathematical preliminaries

### 1.1 Solid angles

Suppose that you wanted to define the word “angle” to someone who had never heard this word. An angle measures how “open” the space is between two lines in a plane. One way you could define an angle more precisely would be based on the geometry of a circle: Imagine drawing two lines from a point, and a circle of radius  $r$  centered on that point. The angle between the two lines,  $\theta$ , is the ratio of the arclength of the circle contained between the two lines to the radius of the circle (figure 1):

$$\theta = \frac{L}{r}.$$

This gives an angle in the usual radian measure, between 0 and  $2\pi$ .

When we go to three-dimensional geometry, we may want to quantify how “open” a cone is. A cone, in this context, is a surface which has a point at one end and extends away from this point. We can imagine very narrow cones and very wide cones. The size of the opening is, roughly, what the solid angle describes. Specifically, we measure the area of intersection between our cone and a sphere of radius  $r$  (figure 2). The solid angle  $\Omega$  is then defined as the ratio

$$\Omega = \frac{S}{r^2}.$$

This formula gives the solid angle in units of **steradians**. The solid angle can take any value between 0 and  $4\pi$  steradians.

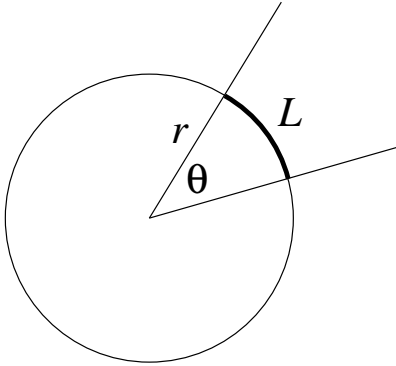


Figure 1: The angle  $\theta$  can be defined as the ratio of  $L$  to  $r$ .

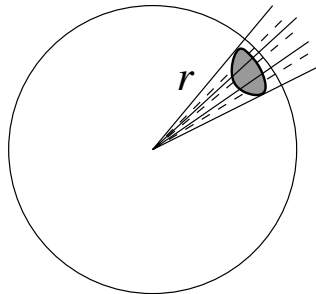


Figure 2: A cone intersects a sphere of radius  $r$  in an area  $S$  (shaded). The solid angle is the ratio of  $S$  to  $r^2$ .

## 1.2 Taylor series

Suppose that  $f(x)$  is a function we want to approximate by a simpler form near some point  $x = x_0$ . For normal functions (continuous, with continuous derivatives), we can use a Taylor series, which fits a polynomial to our function, derivative-by-derivative:

$$f(x) \approx f(x_0) + \left. \frac{df}{dx} \right|_{x=x_0} (x - x_0) + \frac{1}{2} \left. \frac{d^2f}{dx^2} \right|_{x=x_0} (x - x_0)^2 + \dots + \frac{1}{n!} \left. \frac{d^n f}{dx^n} \right|_{x=x_0} (x - x_0)^n.$$

Often, we use Taylor series when  $x - x_0$  is very small. In that case, we often just keep the first non-zero term in the Taylor expansion.

## 2 The freely jointed polymer

We are going to study a very crude model of a polymer in solution, namely the *freely jointed polymer*. In this model, we treat a polymer as being made up of point-like monomers connected by rigid bonds of length  $\ell$ . Both the bond and dihedral angles are unrestricted, i.e. they can assume any value between 0 and  $2\pi$ . This last assumption isn't very realistic, but it turns out that we get reasonably accurate predictions of some polymer properties from this model anyway. Figure 3 shows an example of a polymer generated (in two dimensions) by this model. We will concentrate on *unbranched* polymers, i.e. polymers which consist of a simple string of monomers.

### 2.1 The extent of the polymer along a coordinate axis

We will first calculate the statistical properties of the end-to-end distance along a particular coordinate axis, say the  $x$  axis. In other words, we want the probability distribution of  $x = x_n - x_1$ , where  $x_i$  is the  $x$  coordinate of the  $i$ 'th monomer, and  $n$  is the total number of monomers in the chain. Define  $\psi$  as the angle a bond makes to the  $x$  axis. If we think of the bond as a vector, the  $x$  component of the bond is  $\ell_x = \ell \cos \psi$ . Therefore, if we knew  $p(\psi)$ , the probability density for the angle  $\psi$ , then we could calculate the statistical properties of  $\ell_x$ , and from there the statistical properties of  $x$ .

The probability density is defined such that  $p(\psi) d\psi$  is the probability that some particular measurement  $\Psi$  of this angle satisfies  $\Psi \in (\psi, \psi + d\psi)$ . The region defined by this relationship is a cone whose inner and outer walls are separated by an angle  $\psi$  (figure 4). Now imagine the intersection of this cone with a sphere of radius  $r$ . Since the vectors are randomly oriented, the probability  $p(\psi) d\psi$  is just the fraction of the sphere occupied by the intersection, i.e. the solid angle divided by  $4\pi$ , the total solid angle of a sphere. Specifically, we want the area of the annulus at distance  $r$  from the origin between angles  $\psi$  and  $\psi + d\psi$ . Because  $d\psi$  is very small, we can calculate this area as the circumference of the circle of radius  $r \sin \psi$ , which is  $2\pi r \sin \psi$ , multiplied by the width of the annulus,  $r d\psi$ . Thus the solid angle is  $2\pi r^2 \sin \psi d\psi / r^2 = 2\pi \sin \psi d\psi$ . The probability that  $\Psi$  is between  $\psi$  and  $\psi + d\psi$  is therefore

$$p(\psi) d\psi = \frac{1}{4\pi} 2\pi \sin \psi d\psi = \frac{1}{2} \sin \psi d\psi.$$

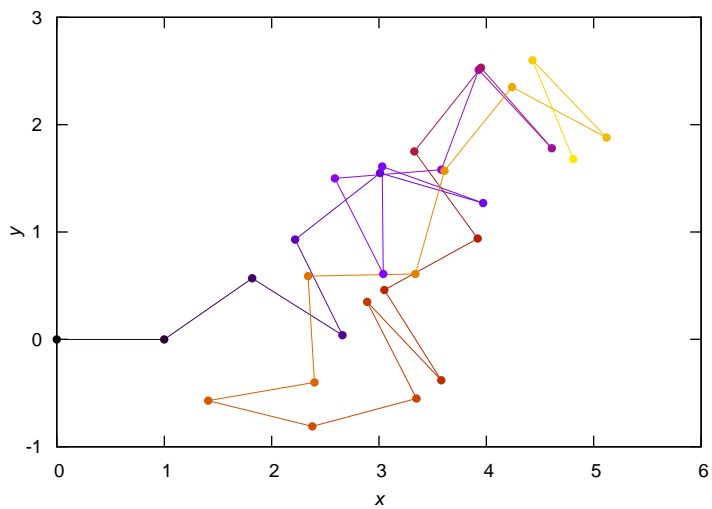


Figure 3: A typical, randomly generated 30-monomer freely jointed polymer with bond length  $\ell = 1$ . The polymer is colored blue at one end and yellow at the other to help you follow the links.

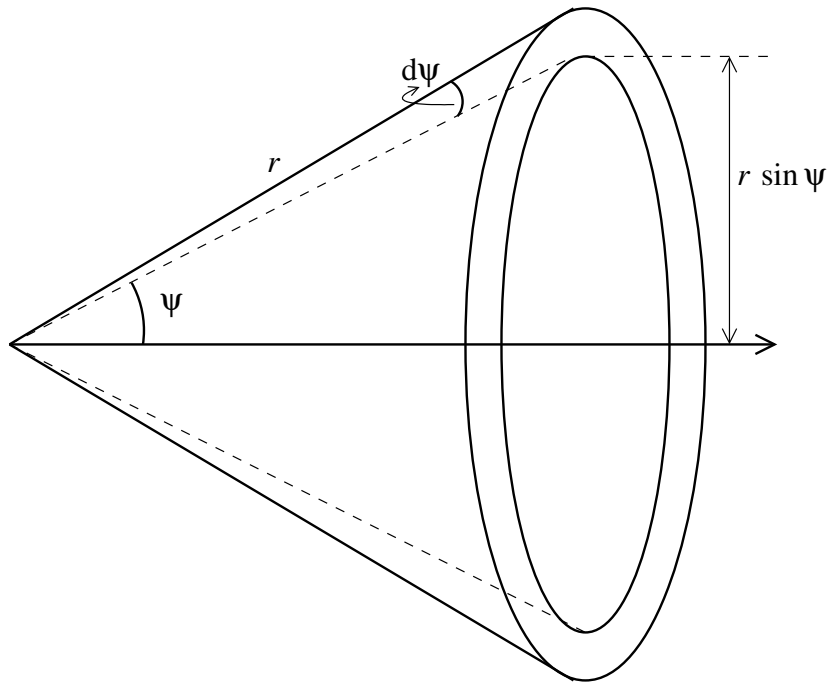


Figure 4: The annular region at distance  $r$  from the origin between angles  $\psi$  and  $\psi + d\psi$  is the one whose area is used to calculate the solid angle, and thence the probability density for angle  $\psi$ .

We can now calculate statistical properties of  $\ell_x$ . Let's start with the average:

$$\begin{aligned}\langle \ell_x \rangle &= \int_0^\pi \ell_x(\psi) p(\psi) d\psi \\ &= \int_0^\pi \ell \cos \psi \cdot \frac{1}{2} \sin \psi d\psi = 0.\end{aligned}$$

This shouldn't be a great surprise: Any given bond is just as likely to go to the left as to the right.

We can also calculate the root-mean-squared displacement along the  $x$  axis:

$$\begin{aligned}\langle \ell_x^2 \rangle &= \int_0^\pi [\ell_x(\psi)]^2 p(\psi) d\psi \\ &= \int_0^\pi (\ell \cos \psi)^2 \cdot \frac{1}{2} \sin \psi d\psi \\ &= \frac{\ell^2}{3}.\end{aligned}$$

$$\therefore \sqrt{\langle \ell_x^2 \rangle} = \frac{\ell}{\sqrt{3}}.$$

This is a measure of the typical distance along the  $x$  axis between one end of a bond and the other.

Now suppose that we want to know the distribution of distances  $x$  between one end of the polymer and the other. If we have a large polymer, then we can assume that, on average, each bond adds to or subtracts from the  $x$  coordinate of the end monomer an amount  $\sqrt{\langle \ell_x^2 \rangle}$ . Let  $n_+$  be the number of monomers that make a positive contribute to  $x$ , and  $n_-$  be the number of monomers that make a negative contribution to  $x$ . Then

$$x = (n_+ - n_-) \sqrt{\langle \ell_x^2 \rangle} = (n_+ - n_-) \frac{\ell}{\sqrt{3}}.$$

Since the sign of the  $x$  displacement is randomly chosen from the two possibilities, this is just a coin-tossing process. The probability of obtaining exactly  $n_+$  heads out of  $n$  tosses is

$$p(n_+|n) = \frac{n!}{2^n n_+! n_-!}.$$

Define

$$m = n_+ - n_-$$

so that

$$x = m \frac{\ell}{\sqrt{3}}. \tag{1}$$

We can rephrase our probability distribution in terms of  $n$  and  $m$ :

$$\begin{aligned}n_+ &= \frac{1}{2}(n + m), \\ \text{and } n_- &= \frac{1}{2}(n - m).\end{aligned}$$

$$\therefore p(m|n) = \frac{n!}{2^n \left[\frac{1}{2}(n + m)\right]! \left[\frac{1}{2}(n - m)\right]!}.$$

Now we take a logarithm of  $p(m|n)$  and apply Stirling's approximation. We are interested in the case where  $n$  is large. We expect that the overwhelmingly largest number of configurations will have  $m \ll n$ , so all the factorials appearing in our probability are factorials of large numbers.

$$\begin{aligned}
\ln p(m|n) &= \ln n! - n \ln 2 - \ln \left[ \frac{1}{2}(n+m) \right]! - \ln \left[ \frac{1}{2}(n-m) \right]. \\
&\approx n \ln n - n - n \ln 2 - \frac{1}{2}(n+m) \ln \left[ \frac{1}{2}(n+m) \right] + \frac{1}{2}(n+m) \\
&\quad - \frac{1}{2}(n-m) \ln \left[ \frac{1}{2}(n-m) \right] + \frac{1}{2}(n-m) \\
&= n \ln n - n \ln 2 - \frac{1}{2}(n+m) \ln \left[ \frac{1}{2}(n+m) \right] - \frac{1}{2}(n-m) \ln \left[ \frac{1}{2}(n-m) \right] \\
&= n \ln n - n \ln 2 - \frac{n}{2} \ln \left[ \left( \frac{1}{2} \right)^2 (n+m)(n-m) \right] + \frac{m}{2} \ln \left( \frac{n-m}{n+m} \right) \\
&= n \ln n - \frac{n}{2} \ln(n^2 - m^2) + \frac{m}{2} \ln \left( \frac{n-m}{n+m} \right) \\
&= n \left\{ \ln n - \frac{1}{2} \ln \left[ n^2 \left( 1 - \left( \frac{m}{n} \right)^2 \right) \right] + \frac{m}{2n} \ln \left( \frac{1 - \frac{m}{n}}{1 + \frac{m}{n}} \right) \right\} \\
&= n \left\{ \frac{m}{2n} \ln \left( \frac{1 - \frac{m}{n}}{1 + \frac{m}{n}} \right) - \frac{1}{2} \ln \left[ 1 - \left( \frac{m}{n} \right)^2 \right] \right\}.
\end{aligned}$$

Recall that  $m/n$  is usually small. We can apply Taylor series to the two logarithms in this expression. Keeping only the first non-zero terms, we get

$$\begin{aligned}
\ln \left( \frac{1 - \frac{m}{n}}{1 + \frac{m}{n}} \right) &\approx -2 \frac{m}{n}, \\
\text{and } \ln \left[ 1 - \left( \frac{m}{n} \right)^2 \right] &\approx - \left( \frac{m}{n} \right)^2. \\
\therefore \ln p(m|n) &\approx n \left\{ - \left( \frac{m}{n} \right)^2 + \frac{1}{2} \left( \frac{m}{n} \right)^2 \right\} = - \frac{m^2}{2n}.
\end{aligned}$$

Stirling's approximation, in the form we use in this course,<sup>1</sup> leaves out some constants which are negligible at large  $n$ . As a result, we have lost the normalization of the distribution  $p(m|n)$ . The best we can say at this time is that  $p(m|n)$  is proportional to  $e^{-m^2/(2n)}$ . However, that is enough for our purposes. We can use equation 1 to eliminate  $m$  in favor of  $x$ . This gives us the unnormalized probability distribution

$$f(x) = e^{-3x^2/(2n\ell^2)}.$$

---

<sup>1</sup>There are other versions of Stirling's approximation. They are all equivalent for very large values of  $n$ , but some are more accurate at smaller  $n$ .

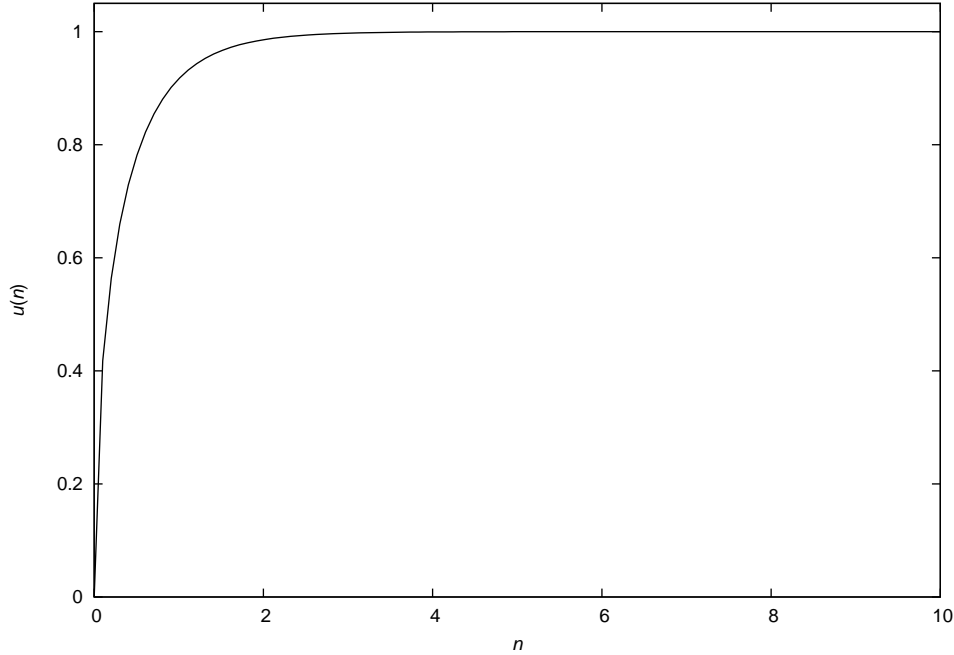


Figure 5: Graph of the function  $u(n) = \operatorname{erf} \sqrt{3n/2}$ .

If we always took steps of exactly  $\ell/\sqrt{3}$  along the  $x$  axis, that would be the end of the story, give or take normalizing this distribution. However, this is only an approximation valid for averaging over a large number of steps. In fact,  $x$  is a continuous variable, not a discrete one like  $m$ . We should therefore interpret  $f(x)$  as a *probability density*. The normalization constant  $C$  is found as follows:

$$\begin{aligned} C &= \left( \int_{-n\ell}^{n\ell} f(x) dx \right)^{-1} \\ &= \frac{1}{\ell \operatorname{erf} \left( \sqrt{3n/2} \right)} \sqrt{\frac{3}{2\pi n}}. \end{aligned}$$

Note that the limits of integration are set by the maximum possible extent of the polymer, which we obtain only if every single bond points in the same direction, either along the  $+x$  or  $-x$  semi-axis. The function  $\operatorname{erf}(\cdot)$  is the error function. It often comes up when computing integrals involving  $e^{-kx^2}$ . We don't need to know much about this function since programs like Maple and Excel know how to calculate its value. Figure 5 shows the function  $u(n) = \operatorname{erf} \sqrt{3n/2}$ . Notice how fast this function approaches  $u = 1$ . Since we are mostly interested in longer polymer chains, we can therefore set  $u(n) = 1$  in our normalization factor. The probability distribution becomes

$$p(x) dx = \frac{1}{\ell} \sqrt{\frac{3}{2\pi n}} e^{-3x^2/(2n\ell^2)} dx. \quad (2)$$



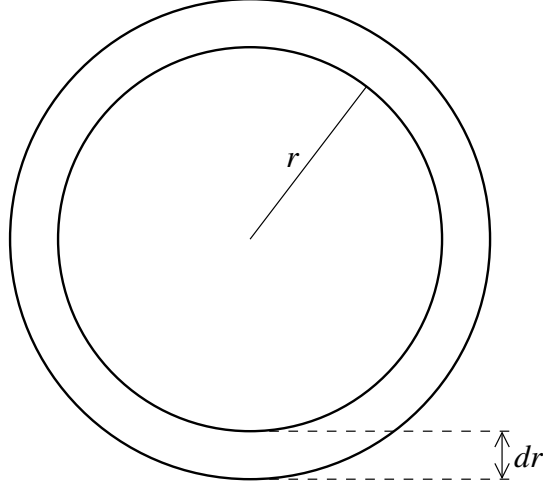


Figure 6: Thin spherical shell used to determine the radial probability density.

Note that this is a Gaussian distribution. It is normalized provided we extend the range of integration to be  $(-\infty, \infty)$ , which of course is consistent with assuming that  $n$  is large.

## 2.2 End-to-end distance in three dimensions

Equation 2 applies to any direction in space. We therefore instantly know the probability density for the three independent Cartesian axes:

$$\begin{aligned} p(x, y, z) dx dy dz &= [p(x) dx] [p(y) dy] [p(z) dz] \\ &= \left( \frac{1}{\ell} \sqrt{\frac{3}{2\pi n}} \right)^3 e^{-3r^2/(2n\ell^2)} dx dy dz, \end{aligned}$$

$$\text{where } r^2 = x^2 + y^2 + z^2.$$

$p(x, y, z) dx dy dz$  is the probability that the relative position of the last monomer in the polymer relative to the first is in the small box of edge lengths  $dx$ ,  $dy$  and  $dz$  with one corner at  $(x, y, z)$ . Of course, we don't usually care how the polymer is oriented. What we usually want to know is how far the two ends are from each other. In other words, we want  $p(r) dr$ . This would be the probability that the relative position vector  $\mathbf{r} = (x_n - x_1, y_n - y_1, z_n - z_1)$  has a length between  $r$  and  $r + dr$ , irrespective of the angle this vector makes to the Cartesian axes. The vectors which satisfy this condition define a spherical shell of inner radius  $r$  and thickness  $dr$ . See figure 6. What we need to do is to integrate (add up)  $p(x, y, z) dx dy dz$  over the interior of the shell. Since  $p(x, y, z)$  only depends on  $r$ , it is essentially constant in our thin shell, and we can pull it out of the integral. All that is left to do is to integrate  $dx dy dz$  over our shell, but this is just the volume of the shell, i.e.  $4\pi r^2 dr$  (area times thickness). The desired probability distribution is therefore

$$p(r) dr = \left( \frac{1}{\ell} \sqrt{\frac{3}{2\pi n}} \right)^3 e^{-3r^2/(2n\ell^2)} \cdot 4\pi r^2 dr.$$

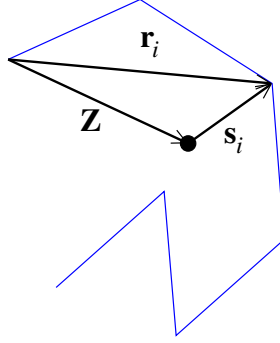


Figure 7: Vectors used to calculate the radius of gyration. The polymer is shown in blue. The heavy dot represents the centre of mass.

Note that there really isn't anything special about the first and last monomers in a chain. Provided  $n$  isn't too small, this equation therefore gives the distribution of distances between any two monomers separated by  $n$  bonds in the chain.

### Exercises

1. Calculate the mean distance between the two ends of a polymer assuming the freely jointed chain model.
2. Calculate the standard deviation of the distance between the two ends of a freely jointed polymer.

## 2.3 Radius of gyration

Many polymer properties depend on the **radius of gyration**, which is the root-mean-squared distance of the monomers from the centre of mass of the polymer. The radius of gyration turns out to be related to the root-mean-squared distance between two monomers, so we start by computing this quantity:

$$\begin{aligned} \langle r^2 \rangle &= \int_0^\infty r^2 p(r) dr = n\ell^2. \\ \therefore \sqrt{\langle r^2 \rangle} &= \ell\sqrt{n}. \end{aligned} \quad (3)$$

Again note that this is the rms distance between any two monomers separated by  $n$  bonds.

We need to define a number of vectors:

- $\mathbf{r}_i$  is the vector from one end of the chain (i.e. monomer 1) to monomer  $i$ .
- $\mathbf{s}_i$  is the vector from the centre of mass to monomer  $i$ .
- $\mathbf{Z}$  is the vector from monomer 1 to the centre of mass. ( $\mathbf{Z} = -\mathbf{s}_1$ .)

Figure 7 illustrates these vectors. From these definitions, we have

$$\mathbf{s}_i = \mathbf{r}_i - \mathbf{Z}. \quad (4)$$

The centre of mass is defined by  $\sum_i s_i = 0$ . Summing equation 4 over all monomers, we get

$$0 = \sum_i \mathbf{r}_i - n\mathbf{Z},$$

$$\text{or } \mathbf{Z} = \frac{1}{n} \sum_i \mathbf{r}_i. \quad (5)$$

$$\therefore Z^2 = \mathbf{Z} \cdot \mathbf{Z} = \frac{1}{n^2} \sum_i \sum_j \mathbf{r}_i \cdot \mathbf{r}_j. \quad (6)$$

The radius of gyration is, as mentioned above, the rms distance from the centre of mass or, in terms of the vectors we have just defined, the rms value of  $s_i$ . We start by calculating the squared sum of the distances from the centre of mass:

$$\begin{aligned} \sum_i s_i^2 &= \sum_i s_i^2 = \sum_i \mathbf{s}_i \cdot \mathbf{s}_i \\ &= \sum_i (\mathbf{r}_i - \mathbf{Z}) \cdot (\mathbf{r}_i - \mathbf{Z}). \end{aligned}$$

This last relationship was obtained using equation 4. Thus we have

$$\sum_i s_i^2 = \sum_i r_i^2 + nZ^2 - 2 \sum_i \mathbf{Z} \cdot \mathbf{r}_i. \quad (7)$$

Let's work on the last term in this equation:

$$\sum_i \mathbf{Z} \cdot \mathbf{r}_i = \mathbf{Z} \cdot \sum_i \mathbf{r}_i = \mathbf{Z} \cdot n\mathbf{Z} = nZ^2,$$

where we used equation 5 to eliminate the sum. Going back now to equation 7, we have

$$\sum_i s_i^2 = \sum_i r_i^2 - nZ^2.$$

Now we use equation 6:

$$\sum_i s_i^2 = \sum_i r_i^2 - \frac{1}{n} \sum_i \sum_j \mathbf{r}_i \cdot \mathbf{r}_j. \quad (8)$$

The dot product is defined by

$$\mathbf{r}_i \cdot \mathbf{r}_j = r_i r_j \cos \theta,$$

but the cosine rule says that

$$r_i^2 + r_j^2 - 2r_i r_j \cos \theta = r_{ij}^2,$$

where  $r_{ij}$  is the distance between monomers  $i$  and  $j$ . Combining these two equations to eliminate  $r_i r_j \cos \theta$ , we get

$$\mathbf{r}_i \cdot \mathbf{r}_j = \frac{1}{2} (r_i^2 + r_j^2 - r_{ij}^2).$$

Substituting this relationship into equation 8, we get

$$\begin{aligned}
\sum_i s_i^2 &= \sum_i r_i^2 - \frac{1}{2n} \sum_i \sum_j (r_i^2 + r_j^2 - r_{ij}^2) \\
&= \sum_i r_i^2 - \frac{1}{2n} \left( \sum_j \sum_i r_i^2 + \sum_i \sum_j r_j^2 - \sum_i \sum_j r_{ij}^2 \right) \\
&= \sum_i r_i^2 - \frac{1}{2n} \left( n \sum_i r_i^2 + n \sum_j r_j^2 - \sum_i \sum_j r_{ij}^2 \right) \\
&= \sum_i r_i^2 - \frac{1}{2n} \left( 2n \sum_i r_i^2 - \sum_i \sum_j r_{ij}^2 \right) \\
&= \frac{1}{2n} \sum_i \sum_j r_{ij}^2 \\
&= \frac{1}{n} \sum_j \sum_{i<j} r_{ij}^2.
\end{aligned}$$

The last transformation uses the fact that each of the pairwise distances occurs twice in the previous sum. To get the radius of gyration, we first need to get the average of  $s^2$ :

$$\langle s^2 \rangle = \frac{1}{n} \sum_i \langle s_i^2 \rangle = \frac{1}{n^2} \sum_j \sum_{i<j} \langle r_{ij}^2 \rangle.$$

For a freely jointed polymer,  $\langle r_{ij}^2 \rangle$  is given by equation 3 with  $n = j - i$ . We therefore have

$$\langle s^2 \rangle = \frac{\ell^2}{n^2} \sum_j \sum_{i<j} (j - i) = \frac{\ell^2}{n^2} \sum_j \sum_{i=1}^{j-1} i.$$

The last equality is just a reordering of the inner sum, as you can see for yourself by writing down the terms in the two forms of the inner sum. The sums we need to calculate are well known. I just looked them up in a reference book. We will also make a few approximations using the fact that we are interested in large polymers, i.e. polymers with a large  $n$ . First,  $\sum_{i=1}^{j-1} i = j(j-1)/2$ . We will then sum this quantity over values of  $j$  from 1 to  $n$ . Of course, most of these values will be large, so we can approximate this sum as  $j^2/2$ . This leaves us with the following sum:  $\sum_{j=1}^n j^2 = n(n+1)(2n+1)/6 \approx n^3/3$ . Putting it all together, we get

$$\langle s^2 \rangle = \frac{\ell^2}{n^2} \frac{1}{2} \frac{n^3}{3} = \frac{n\ell^2}{6}.$$

Now we use equation 3 again, obtaining the result

$$\begin{aligned}
\langle s^2 \rangle &= \langle r^2 \rangle / 6, \\
\text{or } \sqrt{\langle s^2 \rangle} &= \sqrt{\langle r^2 \rangle} / 6.
\end{aligned}$$

This quantity is the radius of gyration we wanted to calculate. Note that the radius of gyration for a *linear* polymer—this isn't true for branched polymers—is directly proportional to the average end-to-end distance.

What good is the crude theory developed in these pages? It turns out that several predictions of this model—the Gaussian distribution of displacements along a coordinate axis, the form of the probability density for the end-to-end distance, and the relationship between the radius of gyration and root mean squared end-to-end distance—are all roughly independent of the model details. Even if we added realistic bond and dihedral angle constraints, we would get about the same results.

### Exercises

1. A linear polymer has monomers which are about  $15 \text{ \AA}$  in length. Plot the radius of gyration as a function of  $n$ , for  $n$  up to 1000.