

## A one-dimensional random walk model for polymer chain.

Consider a chain consisting of  $N$  segments each of length 1. Assume that the  $x$ -axis is in the direction of the chain, and let us place the beginning of the chain in the origin. We are interested in finding the probability that the end of the chain will be at a distance  $L$  from the origin.  $L$  is also called the “end-to-end” distance.

This probability equals the number of possible ways to arrange the chain so that its end is at the distance  $L$  (let’s call it  $\Omega(L,N)$ ) *divided* by the total number of all possible arrangements for a chain of length  $N$  ( $\Omega(N)$ ) {here I use  $\Omega$  instead of  $W$  for the weight, to be consistent with the textbook}.

$$P(L,N) = \Omega(L,N)/\Omega(N) \quad (1)$$

In a 1D chain, each segment can be considered as a vector of length 1 pointing in either positive (+) or negative (–) direction along the  $x$ -axis. If  $N_+$  is the number of segments oriented in the “+” direction and  $N_-$  is the number of segments in the “–” direction for a given conformation of the chain, then the end-to-end distance for this conformation is

$$L = N_+ - N_- \quad (2)$$

Actually, the way it is defined here,  $L$  gives the position (along the  $x$ -axis) of the end of the chain, given the beginning is in the origin. So,  $L$  has the meaning of the  $x$ -coordinate of the end-to-end vector.

Note that the total number of segments equals  $N$ :

$$N_+ + N_- = N. \quad (3)$$

From Equation (3) we obtain  $N_- = N - N_+$ , and substituting it for  $N_-$  in Eq.(2) we get:

$$L = 2N_+ - N = 2(N_+ - N/2) \quad (4)$$

This means that for a given  $N$ ,  $L$  is directly determined by the number of + segments in the chain. Therefore,  $\Omega(L,N)$  can be replaced by  $\Omega(N_+, N)$ , i.e. the number of possible ways to arrange  $N_+$  plus-segments in the  $N$ -segments chain:

$$\Omega(L,N) = \Omega(N_+, N) \quad (5)$$

We assume that for each segment there is an equal probability to be oriented in the “+” and in the “–” direction, *independent* of any other segment in the chain (a random-walk model). Naturally, we can expect that for the most probable conformation,

$N_+ = N_- = \frac{1}{2} N$  (assume  $N$  is a big number).

For convenience, we can label each segment by a + or a – depending on its orientation, so that each conformation of the chain is characterized by a string of alternating pluses and minuses, the total number of positions being  $N$ ). For example, a chain of 5 segments with  $N_+=2$  and  $N_-=3$  can be obtained if the following 10 ways:

( + + - - - )      ( - + + - - )      ( - - + + - )      ( - - - + + )  
 ( + - + - - )      ( - + - + - )      ( - - + - + )  
 ( + - - + - )      ( - + - - + )  
 ( + - - - + )

(The relative arrangement of these 10 possible conformations does not matter: I grouped them just for convenience of keeping track of where I put the pluses).

Then counting of all possible ways to have a chain with  $N_+$  segments in the + direction and  $N_-$  segments in the – direction is the same problem as asking how many ways are there to distribute  $N_+$  pluses over  $N$  positions in a string of  $N$  characters. (We don't have to worry about distributing the minuses, because they will fill the non-plus spaces automatically). This is *exactly* the same problem as distributing  $N_+$  balls over  $N$  empty bins such that each bin can be occupied by no more than one ball. We can start with a set of  $N$  empty bins (analog = a string of all minuses). There are  $N$  possibilities to place the first ball. After it is placed (analog=a minus is replaced by a plus), there are now  $N-1$  bins available for placing the second ball. Similarly, after the first two balls were placed, there will remain  $N-2$  possibilities to place the third ball and so on.

Then the total number of possibilities to place  $N_+$  balls, taking them one by one, is a product of the number of possibilities we have for placing each ball, i.e.

$$P_{N,N_+} = N*(N-1)*(N-2)*...*(N-N_++1) = N!/(N-N_+)!$$

In the above mentioned case of  $N=5$  and  $N_+=2$  this gives:  $P_{5,2} = 5!/(5-2)! = 5!/3! = 20$ .

We are not done yet. Because all balls (or pluses) are identical and indistinguishable, we have overestimated the number of possible arrangements. For example, the first conformation, ( + + - - - ), has occurred in our method of counting twice:

as

([first ball],[second ball],[empty],[empty],[empty]) (or ([first +],[second +],[-],[-],[-]))

and as

([second ball],[first ball],[empty],[empty],[empty]) (or ([second +],[first +],[-],[-],[-])).

And therefore it was counted twice. These two conformations are **indistinguishable!**

To correct the number we obtained for the fact that the balls (pluses) are all indistinguishable, we have to divide  $P_{N,N_+}$  by the number of possible permutations for the  $N_+$  objects, which is  $N_+!$ .

So, the correct number of possible (distinguishable) arrangements of  $N_+$  pluses among  $N$  places is

$$\Omega(N_+,N) = \frac{P_{N,N_+}}{N_+!} = \frac{N!}{N_+!(N-N_+)!} = \frac{N!}{N_+!N_-!} \quad (6)$$

In the example above, we get  $\Omega(L,N) = 5!/(2! * 3!) = 10$ .

Note that here pluses are not better than minuses. You can do a similar counting job assuming that you distribute  $N_-$  minuses over a string of  $N$  spaces (initially filled with pluses). You will get the same result. Note that the end-result in Eq(6) is symmetric over  $N_+$  and  $N_-$ , i.e. it will not change if you replace  $N_+$  with  $N_-$  and  $N_-$  with  $N_+$  at the same time.

**Now let's do some math.** Let's first express  $N_+$  and  $N_-$  in terms of  $L$  (in the handout notes they use  $x$  which is  $L/2$ ). From Eq.(4) we get  $N_+ = (N+L)/2$ ; and then from Eq.(3):  $N_- = (N-L)/2$ , so that

$$\Omega(L, N) = \frac{N!}{\left(\frac{N+L}{2}\right)! \left(\frac{N-L}{2}\right)!}. \quad (7)$$

As we did in statistical mechanics, it is convenient to get rid of the factorials by considering a logarithm of  $\Omega$  and taking advantage of Stirling's approximation. Let us use a separate letter/variable for this logarithm (let's call it  $Y$ ):

$$Y = \ln\{\Omega(L, N)\} \quad (8)$$

We can always obtain  $\Omega(L, N)$  from  $Y$  by using an inverse transformation:

$$\Omega(L, N) = e^Y \quad (9)$$

substituting Eq.(7) into Eq.(8) and using Stirling's approximation ( $\ln N! = N \ln N - N$ ) we get:

$$Y = N \ln N - N - \frac{N+L}{2} \ln \frac{N+L}{2} + \frac{N+L}{2} - \frac{N-L}{2} \ln \frac{N-L}{2} + \frac{N-L}{2}$$

(the terms not containing the logarithm cancel!)

$$= N \ln N - \frac{N+L}{2} \ln \frac{N+L}{2} - \frac{N-L}{2} \ln \frac{N-L}{2}$$

( $\ln\{(N+L)/2\} = \ln(N+L) - \ln 2$ , etc.)

$$= N \ln N - \frac{N+L}{2} \ln(N+L) - \frac{N-L}{2} \ln(N-L) + N \ln 2$$

(we actually don't care much about the terms that do not contain  $L$  but let's keep them for now)

$$= N \ln N + N \ln 2 - \frac{N+L}{2} \ln(N+L) - \frac{N-L}{2} \ln(N-L)$$

Here we take advantage of the fact that  $N$  is a big number, and as long as  $L \ll N$ ,  $L/N$  is a small number. Let's call it  $\varepsilon$ , for simplicity:

$$L/N = \varepsilon. \quad (\text{keep in mind that } \varepsilon \ll 1) \quad (10)$$

We will use the following property of the logarithmic function (Taylor series expansion):

$$\ln(1+\varepsilon) = \varepsilon - \varepsilon^2/2 - O(\varepsilon^3).$$

Then  $N+L = N(1+\varepsilon)$ , and  $\ln(N+L) = \ln N + \ln(1+\varepsilon) = \ln N + \varepsilon - \varepsilon^2/2$  (we will keep only terms up to the  $\varepsilon^2$  and ignore higher powers of  $\varepsilon$ , as these are much smaller numbers. Similarly,  $\ln(N-L) = \ln N + \ln(1-\varepsilon) = \ln N - \varepsilon - \varepsilon^2/2$ ).

We then obtain:

$$Y = N \ln N + N \ln 2 - \frac{N}{2} \left[ (1 + \varepsilon) \left( \ln N + \varepsilon - \frac{\varepsilon^2}{2} \right) + (1 - \varepsilon) \left( \ln N - \varepsilon - \frac{\varepsilon^2}{2} \right) \right]$$

The expression in the square brackets can be simplified as follows:

$$[\dots] = \ln N + \varepsilon \ln N + \varepsilon + \varepsilon^2 - \varepsilon^2/2 - \varepsilon^3/2 + \ln N - \varepsilon \ln N - \varepsilon + \varepsilon^2 - \varepsilon^2/2 + \varepsilon^3/2 = 2 \ln N + \varepsilon^2$$

and then we get

$$Y = N \ln N + N \ln 2 - N \ln N - N \varepsilon^2/2 = N \ln 2 - L^2/(2N). \quad (11)$$

The last expression was obtained using the definition of  $\varepsilon$  (Eq.(10)). Recalling Eq.(9) we finally get:

$$\Omega(L, N) = 2^N e^{-\frac{L^2}{2N}} \quad (12)$$

This is the main result that we wanted to obtain:

for a 1D random chain composed of N segments of length 1

**the number of possible conformations having the end-to-end distance L scales with L as  $\exp(-L^2/(2N))$ .**

Then the probability that a randomly picked conformation of such a chain has the end-to-end distance L is (recall Eq.(1))

$$P(L, N) = C_N e^{-\frac{L^2}{2N}} \quad (\text{final result!}) \quad (13)$$

where  $C_N$  is a normalization factor that does not depend on L and can be determined from the condition that the sum of probabilities over all possible states of the chain (i.e. over all possible L

values) should equal 1:  $\sum_L P(L, N) \approx \int P(L, N) dL = 1$ .

For the 1D problem considered here, this condition gives (recall, L is a projection of the end-to-

end vector):  $C_N \sum_{L=-N}^N e^{-\frac{L^2}{2N}} \approx C_N \int_{-\infty}^{\infty} e^{-\frac{L^2}{2N}} dL = C_N \sqrt{2\pi N} = 1$ ; so  $C_N = (2\pi N)^{-1/2}$

-----  
 Note that the total number of all possible conformations (independent of what the end-to-end distance is) of a 1D chain of N segments is  $\Omega(N) = 2^N$ . You can easily obtain this number by realizing that each segment can be in two states: either in the + or in the - direction. If we put this result, together with Eq(12) into Eq(1), we get Eq(13) with  $C_N=1$ . This  $C_N$  is different from the correct normalization: this is due to the Stirling's approximation that we used to derive  $\Omega(L,N)$ : we have ignored terms containing  $N^{-1/2}$ . The most accurate form of the Stirling's approximation is  $N! = N^{N+1/2} e^{-N} (2\pi)^{1/2}$ .

$$\text{Useful table integrals: } \int_0^{\infty} e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a}}; \int_0^{\infty} e^{-ax^2} x^{2n} dx = \sqrt{\frac{\pi}{a}} \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)}{2^{n+1} a^n} \quad (\text{for } n > 0)$$