

# SUMMER RESEARCH PROJECT

DONE AT



ON

---

## PERSISTENT RANDOM WALK

---

UNDER THE GUIDANCE OF

PROF. ABHISHEK DHAR  
INTERNATIONAL CENTRE FOR THEORETICAL SCIENCES

BY

MAHESH GANDIKOTA  
SCHOOL OF PHYSICAL SCIENCES  
NATIONAL INSTITUTE OF SCIENCE EDUCATION AND RESEARCH



REPORT SUBMITTED ON  
OCTOBER 22, 2013

## Abstract

Unbiased, biased and persistent random walks are discussed. The results of the simulations of their basic moments are shown. The models and thermodynamics that is relevant to [2] which is a work related to the behaviour of semi-flexible polymers as predicted by the WLC model is explained in a section. Basic theory and few simulations pertaining to the behaviour of E.Coli forms the last section.

## Contents

<b>1</b>	<b>Unbiased Random Walker</b>	<b>3</b>
1.1	Introduction . . . . .	3
1.2	Derivation of probability density . . . . .	4
1.2.1	Counting paths method . . . . .	4
1.2.2	Difference equation method . . . . .	5
1.3	Simulation . . . . .	7
<b>2</b>	<b>Biased Random Walker</b>	<b>8</b>
2.1	Derivation of probability density . . . . .	8
2.2	Simulation . . . . .	9
<b>3</b>	<b>Persistent Random Walker</b>	<b>10</b>
3.1	Differential equation . . . . .	10
3.2	Simulation . . . . .	11
<b>4</b>	<b>Triple Minima in the Free Energy of the Semiflexible Polymers</b>	<b>12</b>
4.1	Freely jointed model . . . . .	13
4.2	Freely Rotating Chain (FRC) model . . . . .	13
4.3	Kratky-Porod model . . . . .	15
4.4	Worm Like Chain (WLC) model . . . . .	16
4.5	Relevant Thermodynamics . . . . .	17
4.6	Relevance to biology . . . . .	19
<b>5</b>	<b>Run-and-tumble bacteria</b>	<b>19</b>
5.1	Theory . . . . .	20
5.2	Simulation . . . . .	23
<b>6</b>	<b>Acknowledgements</b>	<b>26</b>
<b>7</b>	<b>References</b>	<b>26</b>

# 1 Unbiased Random Walker

## 1.1 Introduction

Say you have a particle. Which obeys to go right by length  $a$  whenever you flip a head in your coin and goes left by  $a$  when you flip a tail. Say you have many such random particles starting from origin. And you have many friends each having adopted a random walker who obeys your friend's coin. As the ceremony of coin flipping by each of your friends begins, after some time, you find that the distribution of the particles is Gaussian about the origin point.

The reasons for such a distribution of the community of random walkers are below:

1. The walks being random.
2. Randomness calls for using probability to mathematically formulate the behaviour. In this case, the probability for the particle being at a given position involves terms like  $n!$ .
3.  $n!$  approaches a Gaussian for large  $n$ .

These will be explained in some detail in this section.

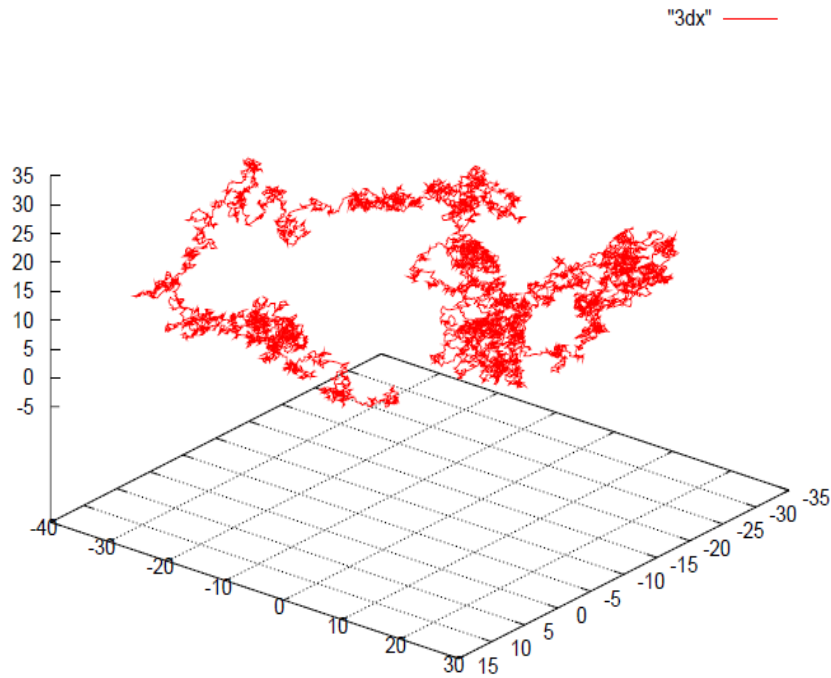


Figure 1: A random walk in a 3D square lattice

## 1.2 Derivation of probability density

### 1.2.1 Counting paths method

Take a random walker on a 1D lattice whose initial position on the lattice is  $X = 0$ . Let him take  $n$  steps in total. He will go right by 1 unit if the tossing of a coin gives a head and left if otherwise. He may be located at any even  $X$  between  $-n$  to  $n$  if  $n$  is even. He will be at an odd  $X$  between  $-n$  to  $n$  if  $n$  is odd.

We are interested in calculating the probability distribution of the particle to be at each of the possible  $X$ . We then would compare the analytical solution got by using probability theory by conducting an experiment (the computer does it).

At each step, the walker can go either right or left, two choices. For  $n$  steps, the possible distinct paths are  $2 * 2 * \dots n$  times which is  $2^n$ .

Let  $P(X, n)$  be the probability of the particle being at  $X$  after  $n$  steps.

$$P(X, n) = \frac{T(X, n)}{2^n} \quad (1)$$

where  $T(X, n)$  are the total number of possible distinct paths of reaching  $X$  after  $n$  steps. More the number of possible paths of ending up at a given  $X$ , more is  $P(X, n)$ .

If  $n_R$  is the total number of right steps in  $n$  steps and  $n_L$  the total number of left steps, then the possible ways of arranging it to give the total distinct paths would be,

$$T(X, n) = \frac{n!}{n_R! n_L!} \quad (2)$$

Now,

$$\begin{aligned} n &= n_R + n_L \\ X &= n_R - n_L \end{aligned} \quad (3)$$

So,

$$\begin{aligned} n_R &= \frac{n + X}{2} \\ n_L &= \frac{n - X}{2} \end{aligned} \quad (4)$$

Substituting in (1), we get

$$P(X, n) = \frac{n!}{\frac{n+X}{2}! \frac{n-X}{2}!} \frac{1}{2^n} \quad (5)$$

If we are intersted in the distribution which is correct only for large  $n$  and small  $X$ , we can use the Stirling's approximation to substitute for all the factorials.

The Stirling's approximation states that for large  $n$ ,

$$n! \approx n^n e^{-n} \sqrt{2\pi n} \quad (6)$$

Using this approximation and substituting it in (5), and making an approximation in the resulting expression which holds true only if  $(X \lesssim O(\sqrt{n}))$ , then we get a neat exponential distribution of the probability.

$$P(X, n) = \left( \frac{2}{\pi n} \right)^{\frac{1}{2}} e^{-\frac{X^2}{2n}} \quad (7)$$

To graduate to a continuum  $x$  from the lattice  $X$ , we define

$$\begin{aligned} x &= X a \\ t &= n \tau \end{aligned} \tag{8}$$

where  $a$  is the step length and  $\tau$  is the time taken for a single step.

The probability density,  $p(x, t)$  is

$$p(x, t) = \frac{P(X, n)}{2a} \tag{9}$$

$P(X, n)$  when plotted versus  $X$  would give a set of discrete non-zero  $P(X, n)$  where each pair of non-zero  $P(X, n)$  has a zero on the y-axis. This is because as we noted that for even  $n$ ,  $P(X, n)$  will be 0 for an odd  $X$ . But in a continuum, we expect a smooth distribution which is differentiable everywhere. To get such distribution, we divide  $P(X, n)$  by  $2a$  instead of  $a$ . Like taking the average of two consecutive steps.

From (7), (8) and (9), we get the probability density

$$p(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \tag{10}$$

where  $D$  is the diffusion constant defined as

$$D = \frac{a^2}{2\tau} \tag{11}$$

It would be clear as to why we have put a 2 in the denominator when we derive the probability distribution using a differential equation.

### Moments

$$\langle X(t) \rangle = a \sum_{i=1}^n \langle \xi_i \rangle = 0$$

where  $\xi_i$  is +1 if it is a right step in  $i$ th step and -1 if it is a left step.

The second moment,

$$\langle X^2(t) \rangle = a^2 \sum_{i,j=1}^n \langle \xi_i \xi_j \rangle = a^2 \left( \sum_{i=1}^n \langle \xi_i^2 \rangle + \sum_{i \neq j}^n \langle \xi_i \xi_j \rangle \right)$$

The second term is zero because no two steps are co-related. So,

$$\langle X^2(t) \rangle = a^2 n \tag{12}$$

### 1.2.2 Difference equation method

We will try to get (10) using a difference equation.

$$P(X, n+1) = \frac{1}{2} [P(X-1, n) + P(X+1, n)] \tag{13}$$

i.e. the probability of the particle being at  $X$  in the  $n+1^{\text{th}}$  step is the sum of the probabilities of the particle being at  $X-1$  in the previous step and moving one step right and the particle being at  $X+1$  and moving one step left.

The above equation can be written as,

$$\frac{P(x, t + \tau) - P(x, t)}{\tau} = D \frac{P(x - a, t) - 2P(x, t) + P(x + a, t)}{a^2}$$

where  $D$  is as defined in (11). We now take limits  $\tau \rightarrow 0$ ,  $a \rightarrow 0$  on both sides such that  $D = \frac{a^2}{2\tau}$  remains constant.

$$\begin{aligned} Lt_{\tau \rightarrow 0} \frac{P(x, t + \tau) - P(x, t)}{\tau} &= D Lt_{a \rightarrow 0} \frac{1}{a} \left[ \frac{p(x + a, t) - p(x, t)}{a} - \frac{p(x, t) - p(x - a, t)}{a} \right] \\ \frac{\partial p(x, t)}{\partial t} &= D Lt_{a \rightarrow 0} \frac{\frac{\partial p(x+a, t)}{\partial x} - \frac{\partial p(x, t)}{\partial x}}{a} \end{aligned}$$

Thus,

$$\frac{\partial p(x, t)}{\partial t} = D \frac{\partial^2 p(x, t)}{\partial x^2} \quad (14)$$

Multiplying the above equation by  $N$ , the total number of walkers, we get

$$\frac{\partial \rho(x, t)}{\partial t} = D \frac{\partial^2 \rho(x, t)}{\partial x^2} \quad (15)$$

where  $\rho(x, t) = Np(x, t)$  is the density of the particles at  $(x, t)$ . This is the *diffusion equation*.

It is now clear from (14) as to why  $D$  was defined with a 2 in the denominator in (11) when we used it first when we were trying to find the probability distribution using counting principles. The reason was just so that in (14), we get such a neat form instead of having a  $2D$  there.

We will now find the solution to (14) using the Fourier transform method.

$$p(x, t) = \int_{-\infty}^{\infty} \tilde{p}(k, t) e^{-ikx} dx$$

Substituting in (14),

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\partial \tilde{p}(k, t)}{\partial t} e^{ikx} dk &= D \int_{-\infty}^{\infty} \tilde{p}(k, t) (-k^2) e^{ikx} dk \\ \Rightarrow \frac{\partial \tilde{p}(k, t)}{\partial t} &= -Dk^2 \tilde{p}(k, t) \\ \Rightarrow \tilde{p}(k, t) &= e^{-Dk^2 t} \tilde{p}(k, 0) \\ \tilde{p}(k, 0) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} p(x, 0) e^{-ikx} dx \\ \text{but } p(x, 0) &= \delta(x) \\ \Rightarrow \tilde{p}(k, 0) &= \frac{1}{2\pi} \\ \Rightarrow \tilde{p}(k, t) &= \frac{1}{2\pi} e^{-Dk^2 t} \end{aligned}$$

Doing the inverse transform,

$$p(x, t) = \int_{-\infty}^{\infty} \tilde{p}(k, t) dk$$

$$\Rightarrow p(x, t) = \int_{-\infty}^{\infty} e^{-Dk^2 t} e^{ikx} dk$$

By change of variables to bring the integral to the form  $\int_{-\infty}^{\infty} e^{-u^2} du$  whose value is  $\sqrt{\pi}$ , we get the solution  $p(x, t)$  to be the same as the expression (10).

The solution to (15) is trivially

$$\rho(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \quad (16)$$

The variance of the density of the community of the random walkers is  $\sqrt{2Dt}$  which increases with time. Thus we see that  $N$  number of random particles, each performing a random walk (not considering their collisions here), diffuses from the initial point where all the particles started their walks.

### 1.3 Simulation

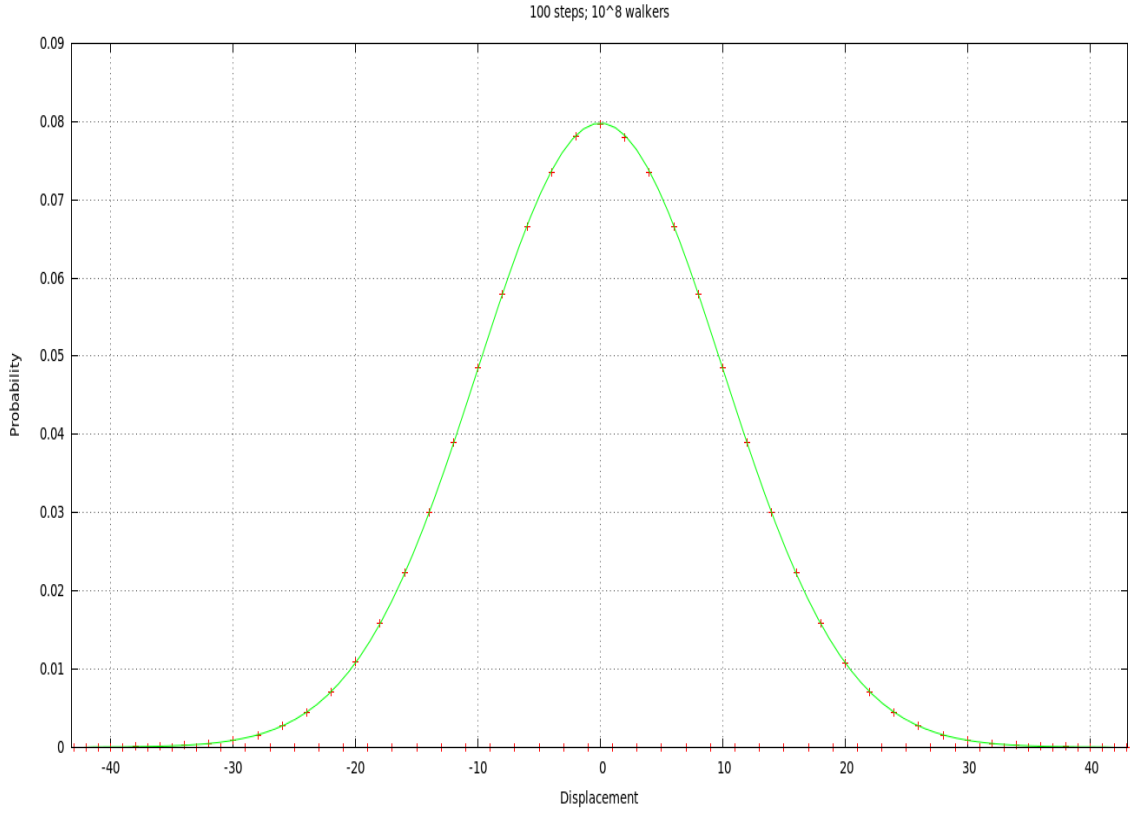


Figure 2: Red squares: Simulation; Green line: Analytical solution

The analytical solution in Fig.(2) is found using equation (7).

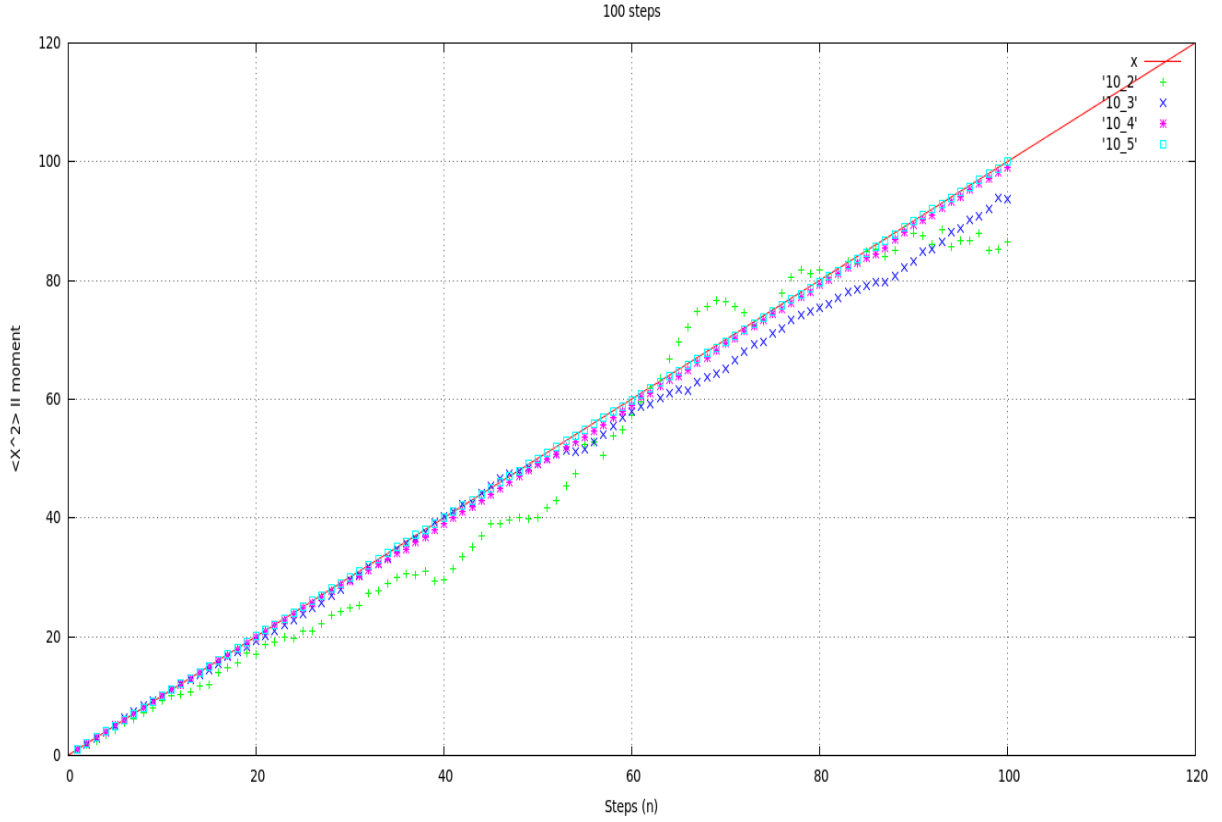


Figure 3: Convergence of simulation to analytical solution (st.line) with increase in number of walkers. No. of walkers is on top-right legend

## 2 Biased Random Walker

Here the random walker has, say more probability of moving to right than to left. Simply put, the coins you use are biased to give more heads than tails. Here too, an analytical solution (for specific cases) can be compared with the simulation.

### 2.1 Derivation of probability density

Here we follow the method of finding the density using the difference equation, graduating to a differential equation and solving it.

Let  $r$  be the probability of going right,  $l$  to go left,  $1-r-l$  to not move at all.

The difference equation is:

$$P(X, n+1) = rP(X-1, n) + lP(X+1, n) + (1-r-l)P(X, n)$$

Going to the continuum case and doing a Taylor expansion taking  $\tau$  and  $a$  to be small,

$$p + \frac{\partial p}{\partial t}\tau = r \left( p - a \frac{\partial p}{\partial x} + \frac{a^2}{2} \frac{\partial^2 p}{\partial x^2} \right) + l \left( p + a \frac{\partial p}{\partial x} + \frac{a^2}{2} \frac{\partial^2 p}{\partial x^2} \right) + (1-l-r)$$

Cancelling some terms and re-arranging

$$\frac{\partial p}{\partial t} = (l-r) \frac{a}{\tau} \frac{\partial p}{\partial x} + (l+r) \frac{a^2}{2\tau} \frac{\partial^2 p}{\partial x^2}$$



When we took  $\tau, a$  to be very small above, i.e. the limits of them tending to zero, we should take the limits such as the constant  $D = (r + l) \frac{a^2}{2\tau}$  remain a constant. Also,

$$v = \frac{a(r - l)}{\tau} = \frac{2D}{a(r + l)}(r - l)$$

should be a constant under the limits on  $\tau, a$ . For the limit of  $v$  to exist, we should also simultaneously take  $Lt(r - l) \rightarrow 0$  such that  $\frac{r-l}{a}$  would remain a constant.

Thus under all such mentioned limits, we get the differential equation which is satisfied by the probability density of a biased random walk.

$$\frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial x^2} - v \frac{\partial p}{\partial x} \quad (17)$$

Note that the above differential equation holds only when  $r$  is almost equal to  $l$ .

## 2.2 Simulation

The below simulation was done by taking  $r + l = 1$ , i.e. the particle always moves either right or left at every step and doesn't stay put at the origin.

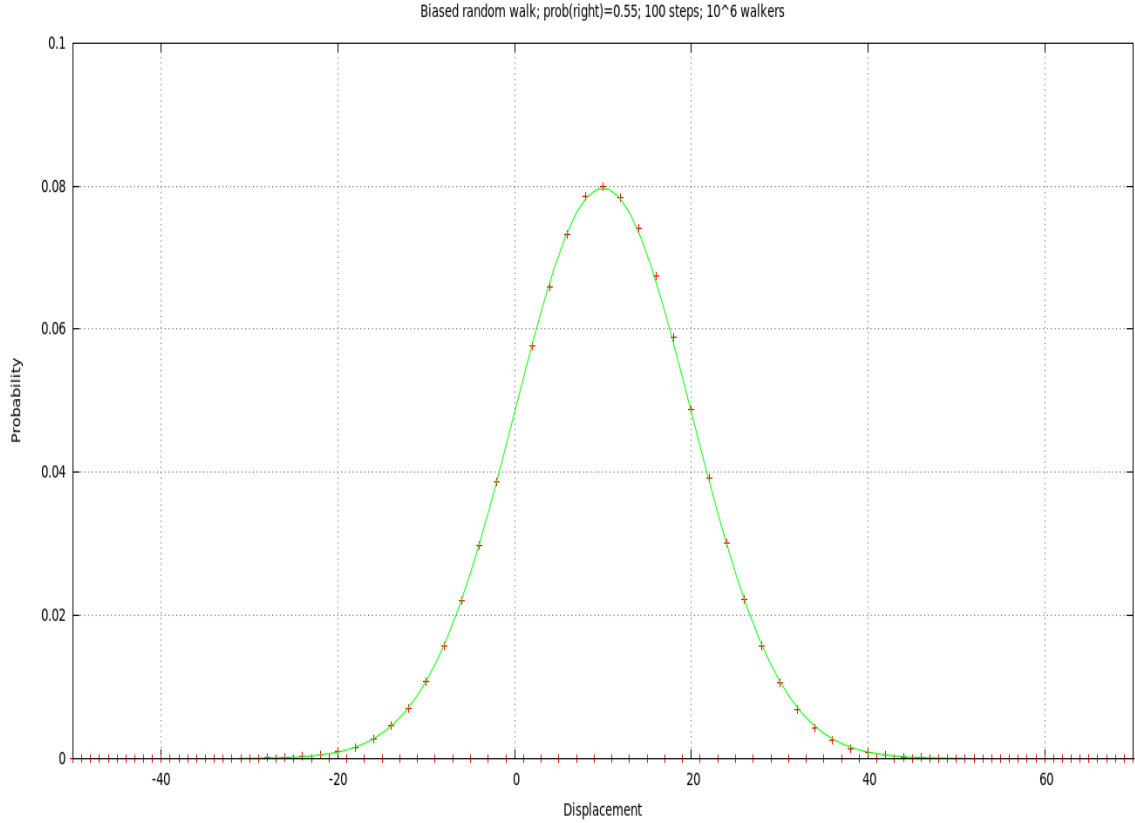


Figure 4: Red squares: Simulation; Green line: Analytical solution

It can be seen that the Gaussian is peaked to the right of 0 and not at 0. The analytical solutions fits to the simulation properly because  $r = 0.55 \approx l = 0.45$ .

### 3 Persistent Random Walker

A persistent walker is someone whose probability of taking the next step in the same direction as the previous is more than going in the opposite direction to the previous step. When the walker takes a step to the right, it is as if with that positive momentum he is more probable to again go right than going left. Likewise, if he has negative momentum, he is more probable to go left again. This is an non-Markovian process. The walker remembers what he did one step before.

#### 3.1 Differential equation

Let  $k$  be the probability that the walker will take the next step in the same direction as the previous. The difference equation in this case is:

$$\begin{aligned} P_+(X, n+1) &= kP_+(X-1, n) + (1-k)P_-(X+1, n) \\ P_-(X, n+1) &= kP_-(X+1, n) + (1-k)P_+(X-1, n) \end{aligned} \quad (18)$$

Here  $P_+(X, n)$  is the probability of the particle to be at  $X$  with positive momentum after  $n$  steps. Similarly  $P_-(X, n)$ .

Subtracting the first equation by  $P_+(X, n)$  on both sides and the second by  $P_-(X, n)$ , going to the continuum of  $X$  and dividing by appropriate constants, we get

$$\begin{aligned} \frac{p_+(x, t+\tau) - p_+(x, t)}{\tau} &= \frac{0.75a}{\tau} \left( \frac{p_+(x-a, t) - p_+(x, t)}{a} \right) + \frac{0.25a}{\tau} \left( \frac{p_-(x+a, t) - p_+(x, t)}{a} \right) \\ \frac{p_-(x, t+\tau) - p_-(x, t)}{\tau} &= \frac{0.75a}{\tau} \left( \frac{p_-(x+a, t) - p_-(x, t)}{a} \right) + \frac{0.25a}{\tau} \left( \frac{p_+(x+a, t) - p_-(x, t)}{a} \right) \end{aligned}$$

By adding and subtracting some terms on the RHS to re-arrange it appropriately and taking limits  $a, \tau$  going to zero on both sides, such that  $G = \frac{a}{\tau}$  remains a constant, we get

$$\begin{aligned} \frac{\partial p_+}{\partial t} &= -kG \frac{\partial p_+}{\partial x} + (1-k)G \frac{\partial p_-}{\partial x} - Lt_{a \rightarrow 0} \frac{1-k}{a} (p_+ - p_-) \\ \frac{\partial p_-}{\partial t} &= kG \frac{\partial p_-}{\partial x} + (1-k)G \frac{\partial p_+}{\partial x} + Lt_{a \rightarrow 0} \frac{1-k}{a} (p_+ - p_-) \end{aligned}$$

Now, we put an additional condition on  $k$  and take limit on both  $k, a$  such that  $\alpha = Lt_{k \rightarrow 1; a \rightarrow 0} \frac{1-k}{a}$  remains a constant.

We then get the coupled differential equations governing the probability density,

$$\begin{aligned} \frac{\partial p_+}{\partial t} &= -G \frac{\partial p_+}{\partial x} - \alpha G (p_+ - p_-) \\ \frac{\partial p_-}{\partial t} &= G \frac{\partial p_-}{\partial x} + \alpha G (p_+ - p_-) \end{aligned} \quad (19)$$

#### II Moment

$$\langle X^2 \rangle = \sum_{i,j=1}^n \langle \xi_i \xi_j \rangle = \sum_i^n \langle \xi_i^2 \rangle + \sum_{i \neq j}^n \langle \xi_i \xi_j \rangle = n + \sum_{i \neq j}^n \langle \xi_i \xi_j \rangle$$

To calculate this, we note the recursion relation:

$$\begin{aligned} \xi_i \xi_j &= 1 \text{ with prob } k; \quad \xi_i \xi_j = -1 \text{ with prob } 1-k \\ \Rightarrow \langle \xi_i \xi_j \rangle &= k \langle \xi_i \xi_{j-1} \rangle - (1-k) \langle \xi_i \xi_{j-1} \rangle = (2k-1) \langle \xi_i \xi_{j-1} \rangle \end{aligned}$$

By carrying on this recursion, we find  $\langle \xi_i \xi_j \rangle = (2k - 1)^{i-j}$ . Summing over all possible  $i, j$  to find the second term of II moment expression, we get the following expression:

$$\langle X^2 \rangle = n + 2[(n-1)(2p-1) + (n-2)(2p-1)^2 + \dots (2p-1)^n]$$

After performing necessary algebra, the closed form of  $\langle X^2 \rangle$  can be seen to be:

$$\langle X^2 \rangle = n + 2 \left[ \left\{ \frac{(n-1)(2p-1)}{2(1-p)} \right\} - \left\{ \frac{(2p-1)^2(1-(2p-1)^{n-1})}{4(1-p)^2} \right\} \right] \quad (20)$$

### 3.2 Simulation

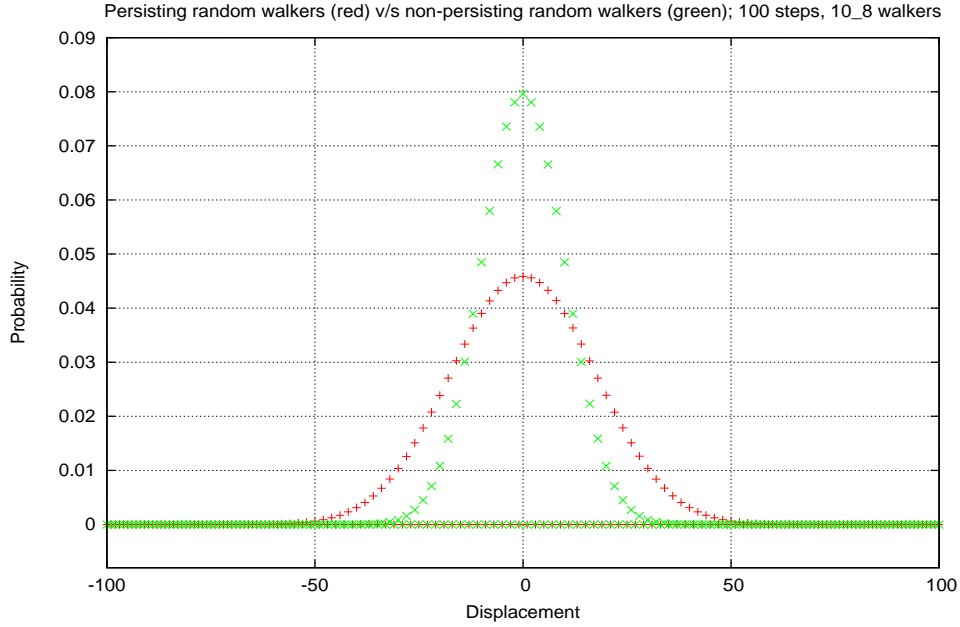


Figure 5: Comparison of probability distributions

It can be seen that the diffusion of the persistent random walkers is more than the unbiased random walkers. This result is intuitive. A persistent random walker can cover more distance due to it's virtue of moving in the same direction for longer times than the unbiased walker.

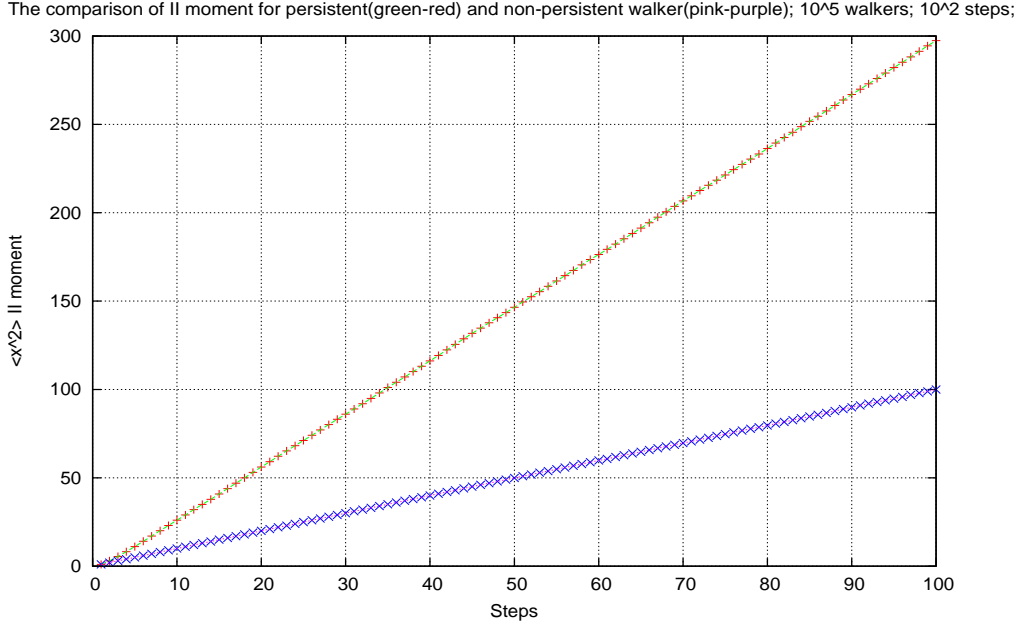


Figure 6: Comparison of II moment; squares: simulation; lines: analytical solution

## 4 Triple Minima in the Free Energy of the Semiflexible Polymers

The intention of this section is not to explain the paper[2] but to supplement it by explaining few of the terminologies/concepts that are to be known to understand the paper. For this purpose, this report has brief descriptions of models of polymers, from the most rudimentary to the WLC model. Also, as the paper assumes familiarity with the thermodynamic potentials, a part of this section provides a brief description of the same.

For a very short description of the paper, I feel the abstract will do better justice than me writing some mess. The abstract of the paper reads thus:

We study the free energy of the worm-like-chain model, in the constant-extension ensemble, as a function of the stiffness  $\lambda$  for finite chains of length  $L$ . We find that the polymer properties obtained in this ensemble are *qualitatively* different from those obtained using constant-force ensembles. In particular, we find that as we change the stiffness parameter,  $t = \frac{L}{\lambda}$ , the polymer makes a transition from the flexible to the rigid phase and there is an intermediate regime of parameter values where the free energy has three minima and both phases are stable. This leads to interesting features in the force-extension curves.

The co-relation of this paper with the former parts of this report is the equivalence of the WLC model with the persistent random walk which has already been discussed.

## 4.1 Freely jointed model

This is the simplest possible model for a polymer. The concept of unbiased random walk is made use of here. However, the walk is not happening over the dimension of time, but over the dimension of space.

Let the polymer be in a 3D space with a co-ordinate frame. Let  $\vec{R}_i$  be the position vector of the  $i^{\text{th}}$  monomer.

$$\hat{u}_n = \frac{\vec{R}_n - \vec{R}_{n-1}}{|\vec{R}_n - \vec{R}_{n-1}|} \quad (21)$$

$\hat{u}_i$  is a unit vector which gives the direction of the  $i^{\text{th}}$  bond and can be called the bond vectors. Let the bond lengths be of unit length.

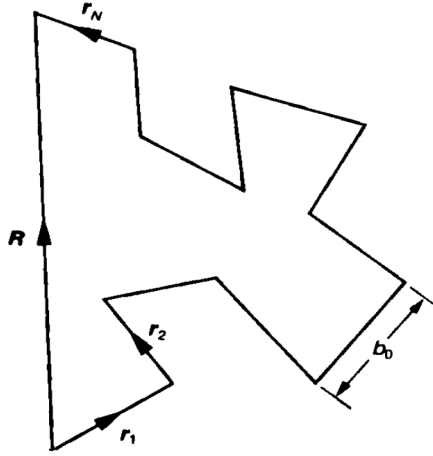


Figure 7: Freely jointed chain[7]. The notations used in this report is  $\hat{u}_i$  instead of  $r_i$  and  $b_0 = 1$ .

There is no co-relation between any two bond vectors as the walk is purely random.

$$\langle \hat{u}_i \cdot \hat{u}_j \rangle = \delta_{ij} \quad (22)$$

This co-relation will be seen to be non-zero for the consecutive models considered here.

To calculate the mean square length of the polymer:

$$\langle \vec{R}^2 \rangle = \left\langle \left( \sum_{j=1}^N \hat{u}_j \right)^2 \right\rangle = \sum_{i=1}^N \langle \hat{u}_i^2 \rangle + \sum_{i \neq j}^N \langle \hat{u}_i \cdot \hat{u}_j \rangle = N$$

Thus the root mean square length of the polymer is proportional to  $\sqrt{N}$ .

## 4.2 Freely Rotating Chain (FRC) model

The angle between consecutive monomers in this model is fixed to be  $\theta$ .

As can be seen in the fig.(8),  $\hat{u}_{j+1}$  has the freedom to be pointing in any direction on the surface of the cone (whose angle with  $\hat{u}_j$  is  $\theta$ ) about  $\hat{u}_j$ .

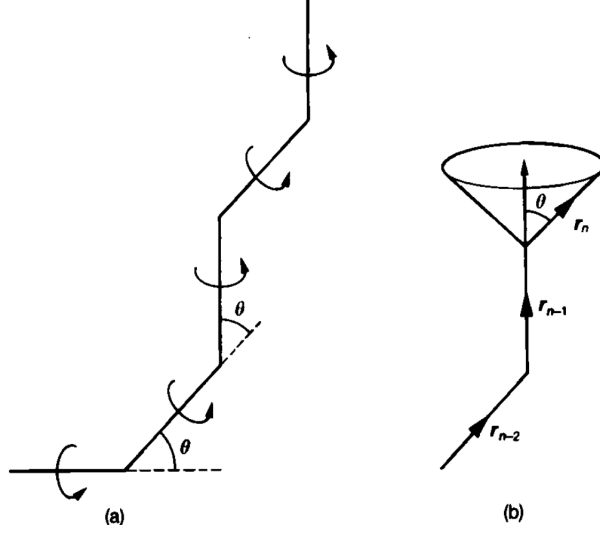


Figure 8: (a) Freely rotating chain (b) The co-relation of  $\mathbf{r}_n$  with  $\mathbf{r}_{n-1}$  is  $\mathbf{r}_{n-1} \cos \theta$ [7].  
The notations used in this report is  $\hat{u}_i$  instead of  $\mathbf{r}_i$ .

Let us feel free to again worry about  $\langle \hat{u}_i \cdot \hat{u}_j \rangle$ . We will calculate this for every model. Let  $i > j$ , evaluating the co-relation between the  $i^{\text{th}}$  and the  $j^{\text{th}}$  bond keeping  $\hat{u}_j, \hat{u}_{j+1}, \dots, \hat{u}_{i-1}$  constant

$$\langle \hat{u}_i \rangle = \cos \theta \hat{u}_{i-1} \quad (23)$$

This is easy to understand because all the components of all possible  $\hat{u}_i$  which are perpendicular to  $\hat{u}_{i-1}$  cancel out on adding, keeping only the component along  $\hat{u}_{i-1}$  alive. Thus,

$$\langle \hat{u}_j \cdot \hat{u}_i \rangle = \cos \theta \langle \hat{u}_j \cdot \hat{u}_{i-1} \rangle \quad (24)$$

Now, that we have a recursion relation, we get

$$\langle \hat{u}_j \cdot \hat{u}_i \rangle = \cos \theta^{|i-j|} = e^{\ln [\cos \theta^{|i-j|}]} = e^{-|i-j| \ln \left[ \frac{1}{\cos \theta} \right]} = e^{-\frac{|i-j|}{\lambda}} \quad (25)$$

where

$$\lambda = \frac{1}{\ln \left[ \frac{1}{\cos \theta} \right]} \quad (26)$$

$\lambda$  is called the *persistence length*. This terminology will be explained in the next model we consider.

From (25), it is clear that the correlation between bonds falls off exponentially with respect to their spatial separation, unlike the previous model where there was no correlation for any pair of bonds.

Now, we can calculate the mean square length of the polymers as calculated from this model.

$$\langle \vec{R}^2 \rangle = \sum_{n=1}^N \sum_{m=1}^N \langle \hat{u}_n \cdot \hat{u}_m \rangle$$

By performing a change of variable,  $m = n + k$ , we rewrite as,

$$\langle \vec{R}^2 \rangle = \sum_{n=1}^N \sum_{k=1-n}^{N-n} \langle \hat{u}_n \cdot \hat{u}_k \rangle$$

Under the long polymer limit,  $\lim_{n \rightarrow \infty}$  and  $\lim_{N-n \rightarrow \infty}$ , we have

$$\begin{aligned} \langle \vec{R}^2 \rangle &= \sum_{n=1}^N \sum_{k=-\infty}^{\infty} \langle \hat{u}_n \cdot \hat{u}_{n+k} \rangle = \sum_{n=1}^N \sum_{k=-\infty}^{\infty} \cos \theta^{|k|} \\ &= \sum_{n=1}^N \left( 1 + 2 \sum_{k=1}^{\infty} \cos \theta^k \right) \\ &= \sum_{n=1}^N [(1 + \cos \theta + \cos^2 \theta + \dots) + \cos \theta (1 + \cos \theta + \cos^2 \theta + \dots)] \end{aligned}$$

Writing the sum of this geometric series,

$$= \sum_{n=1}^N \left[ \frac{1}{1 - \cos \theta} + \cos \theta \frac{1}{1 - \cos \theta} \right] = N \frac{1 + \cos \theta}{1 - \cos \theta}$$

The root mean square length of the polymer is observed to be proportional to  $\sqrt{N}$ .

### 4.3 Kratky-Porod model

The above two models are very simplistic. Kratky and Porod in their 1949 paper[6] considered two more properties of the polymers could be included by considering a Hamiltonian for the polymer of form,

$$H = -K \sum_{j=1}^N \hat{u}_j \cdot \hat{u}_{j+1} = -K \sum_{j=1}^N \cos \theta_j \quad (27)$$

Here,

- The polymer has  $N + 1$  monomers.
- $\hat{u}_j$  is the unit vector in the direction of the  $j^{\text{th}}$  monomer in the 3D space.
- $\theta_j$  is the angle between the  $j^{\text{th}}$  and the  $(j + 1)^{\text{th}}$  monomer.

We see that when consecutive monomers are in the same direction, the energy of the system is lesser and more stable thus. This accounts for the stiffness of the polymers against bond bending. Also, for  $\theta > 90^\circ$ , the value of  $H$  is higher and that configuration is less stable thus. This accounts for the repulsion between the atomic cores of the monomers.

This model puts an *energy* expenditure for different orientations of the bonds. The ensemble of polymers can borrow or dispose energy to a thermal bath. Hence, *temperature* of the bath also comes into the picture.

The concept of temperature and different energies of the polymer for different orientations makes this model distinctly different from the previous two models.

The partition function is,

$$Q_N = \prod_{j=1}^N \int d\Omega_j e^{\beta K \sum_{i=1}^N \cos \theta_i} = (Q_1)^N \quad (28)$$

where

$$Q_1 = 2\pi \int_0^\pi \sin \theta e^{\beta K \cos \theta} d\theta = 4\pi \frac{\sinh \beta K}{\beta K} \quad (29)$$

All the thermodynamic quantities can now be derived from this.

The mean square length of the polymers in the ensemble:

$$\langle \vec{R}^2 \rangle = \left\langle \sum_{j=1}^N \hat{u}_j^2 \right\rangle = \sum_{i,j=1}^N \langle \hat{u}_i \cdot \hat{u}_j \rangle \quad (30)$$

Consider,

$$\langle \hat{u}_i \cdot \hat{u}_j \rangle = \langle \cos \theta \rangle = \frac{\int d\Omega \cos \theta e^{\beta K \cos \theta}}{\int d\Omega e^{\beta K \cos \theta}}$$

On evaluating both the integrals, we get

$$c \equiv \langle \cos \theta \rangle = \frac{\frac{4\pi}{\beta K} \sinh \beta K \left[ \coth \beta K - \frac{1}{\beta K} \right]}{\frac{4\pi}{\beta K} \sinh \beta K} = \coth \beta K - \frac{1}{\beta K}$$

Following the same geometrical argument as given in the FRC model, we have

$$\langle \hat{u}_j \cdot \hat{u}_{j+n} \rangle = c^n = e^{\ln c^n} = e^{-n \ln \frac{1}{c}} = e^{-\frac{n}{\lambda}}$$

where

$$\lambda = \frac{1}{\ln c} \quad (31)$$

is the persistence length. When  $\lambda$  is very small, then the average angle between consecutive angles approaches  $90^\circ$ . Thus the ‘persistence’ of the direction of the previous monomer is lost. If  $\lambda$  is very big, then  $\langle \theta \rangle$  approaches  $0^\circ$ , thus the persistence of the direction of the previous monomer remains. Hence the name ‘persistence length’ for  $\lambda$ .

Now,

$$\langle \vec{R}^2 \rangle = \sum_{i=1}^N \left[ \sum_{j=1}^{i-1} c^{i-j} + 1 + \sum_{j=i+1}^N c^{j-i} \right]$$

Under the long chain approximation,

$$\langle \vec{R}^2 \rangle = \sum_{i=1}^N \left[ \frac{c}{1-c} + 1 + \frac{c}{1-c} \right] = N \frac{1+c}{1-c}$$

It can be observed that even in this model, the root mean square length is proportional to  $\sqrt{N}$ .

#### 4.4 Worm Like Chain (WLC) model

WLC model is got by bringing the Kratky-Porod model to a continuum limit. The Hamiltonian for the polymer is,

$$H = \lambda k_B T \int_0^L \left( \frac{\partial \hat{u}(s)}{\partial s} \right)^2 ds \quad (32)$$

This is the total elastic energy for the given conformation of the polymer. The wriggling of the polymer comes at an energy cost.

The co-relation of tangent vectors:

$$\langle \hat{u}(s) \cdot \hat{u}(s') \rangle = e^{-\frac{|s-s'|}{\lambda}} \quad (33)$$



## 4.5 Relevant Thermodynamics

The thermodynamics concepts used in the paper is given a brief review here.

The first law of thermodynamics is,

$$dU = \bar{d}Q + \bar{d}W \quad (34)$$

$\bar{d}Q$  is the heat flow into the system.

$\bar{d}W$  is the work done on the system.

$\bar{d}$  is written so as to differentiate it from being an exact differential.

$U$  is the internal energy of the system. It is an exact differential in (34) and the sum of two inexact differentials. It is a state function.

For an infinitesimal quasi-static process of a hydrostatic<sup>1</sup> system, (34) may be written as,

$$dU = \bar{d}Q - PdV \quad (35)$$

The **Clausius Theorem** states that for any reversible cycle,

$$\oint_R \frac{\bar{d}Q}{T} = 0 \quad (36)$$

The subscript  $R$  indicates a reversible path. The path independence of this integral shows the existence of a state function, **entropy** defined by,

$$S_f - S_i = \int_R \frac{\bar{d}Q}{T} \quad (37)$$

The infinitesimal change of entropy is,

$$dS = \frac{\bar{d}Q_R}{T} \quad (38)$$

It has to be observed that though the right hand side's numerator is not an exact differential, the left hand side is an exact differential.

(35) may now be written as,

$$dU = -PdV + TdS \quad (39)$$

Other state functions, a complete set of them may be obtained by Legendre differential transformations of the above equation.

- Legendre Transform:  $H = U + PV$  gives,

$$dH = VdP + TdS \quad (40)$$

$H$  is called enthalpy of system.

- Legendre Transform:  $A = U - TS$  gives,

$$dA = -SdT - PdV \quad (41)$$

$A$  is called the Helmholtz free energy of the system.

---

<sup>1</sup>The equilibrium state of the system is completely specified by three thermodynamic co-ordinates:  $P, V, T$

- Legendre Transform:  $G = H - TS$  gives,

$$dG = VdP - SdT \quad (42)$$

$G$  is called the Gibbs free energy of the system.

Notice that the set of all state functions:  $U, H, A, G$  have dimensions of energy.

#### *Conditions for chemical equilibrium*

Consider a hydrostatic system in mechanical and thermal equilibrium, but not in chemical equilibrium<sup>2</sup>. Let the system be in contact with a reservoir of constant temperature  $T$ . Suppose the system undergoes an infinitesimal and irreversible process involving taking  $\bar{d}Q$  heat from the reservoir; let the change in entropy of the system be  $dS$  and the change in entropy of the reservoir be  $dS_o$ . The entropy change of the universe is positive for an irreversible process. So,

$$dS_o + dS > 0 \quad (43)$$

but

$$dS_o = -\frac{\bar{d}Q}{T}$$

Thus by (43),

$$\bar{d}Q - TdS < 0 \quad (44)$$

Using the first law of thermodynamics for the hydrostatic system (35), we get

$$dU + PdV - TdS < 0 \quad (45)$$

1. For constant  $U$  and  $V$ , i.e. for an isolated (because internal energy doesn't change and no work can be done on this hydrostatic system because  $V$  can't be changed) hydrostatic system, the condition for thermodynamic equilibrium is  $dS > 0$ .
2. For constant  $T$  and  $V$ , the condition for thermodynamic equilibrium is  $d(U - TS) < 0$ , i.e.  $dA < 0$ . The equilibrium point is determined by minimizing  $A$ .
3. For constant  $T$  and  $P$ ,  $d(U + PV - TS) < 0$ , i.e.  $dG < 0$ . The equilibrium point is found by minimizing  $G$ .

#### **Co-relation of this section with the paper[2]**

- The thermodynamic properties of the ensemble of polymers whose ends are fixed (volume constant) is obtained from the Helmholtz free energy as in point (2) above.
- The thermodynamic properties of the ensemble of polymers where the force on each polymer is fixed (Pressure constant) is obtained from the Gibbs free energy as in point (3) above.

---

<sup>2</sup>Thermodynamic equilibrium is attained when the system attains mechanical, thermal and chemical equilibrium.

## 4.6 Relevance to biology

Many cellular phenomenon in living cells like active intracellular transport, growth and division of cell require material properties such as recovery of the shape after an external stress has been imposed and dynamic structural reorganization[8]. The structural element of the cell which is responsible for such material properties is the *cytoskeleton* which is a three dimensional assembly of proteins in the eukaryotic cell. The cytoskeleton gives the mechanical and dynamical properties for the cell. And the cytoskeletal polymers (monomers: polypeptides) are semi-flexible polymers at the length scale of less than a few microns. These polymers can take a large number of configurations due to the rotation/bending of chemical bonds. This augmented by the fact that they are in a thermal bath makes it a good idea to study these systems using statistical mechanics.

## 5 Run-and-tumble bacteria

Escherichia Coli is a flagellated microorganism. It is  $1\mu m$  in diameter and  $2\mu m$  in length. These bacteria are motile<sup>3</sup> due to the virtue of their flagella. Their motion comprises of two happenings: *runs* and *tumbles*. During runs, it approximately moves in straight lines for an average time of  $1s$ , at a speed of  $20\mu m s^{-1}$ . During runs, the flagella rotate anti-clockwise. Every run is punctuated by tumbles when the flagella rotate clockwise. They don't exhibit linear motion but change their orientation during this tumble which lasts for about  $0.1s$ . Then, they start their next run in an almost random direction.

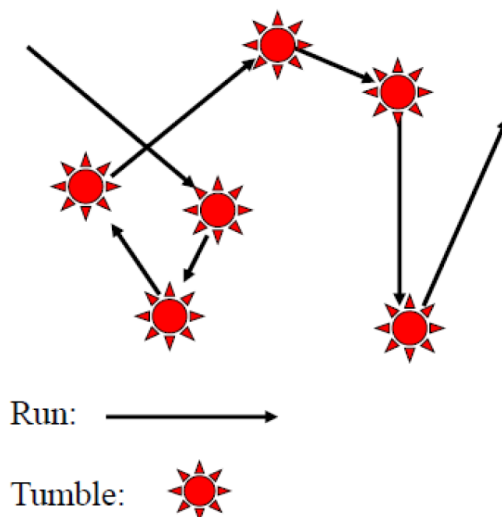


Figure 9: Schematic representation of runs and tumbles in E.Coli[4]

These tumbles can be compared with the collisions in the phenomenon of molecular diffusion. The run lengths are Poisson distributed. The tumbling rate is experimentally seen to be lower in the direction in which the food is increasing and more in the direction in which the food is decreasing. This search for chemical attractants is called *chemotaxis*.

---

<sup>3</sup>Biologist's term for self-propulsion.

It has to be noted that the E.Coli are *still* performing random motion. Their motion is not equivalent to a person who is hungry and moving towards the kitchen. The E.Coli show a behaviour in changing their tumbling rate which will get them to the place with more food with *more* probability.

## 5.1 Theory

**Diffusion constant** The diffusion constant for this stochastic process which obeys Poisson distribution is derived thus:

$$\langle \vec{x}^2 \rangle = \left\langle \sum_{i,j}^N \vec{r}_i \cdot \vec{r}_j \right\rangle$$

Again, making use of the fact that the steps  $\vec{r}_i$  are not co-related, we have

$$\langle \vec{x}^2 \rangle = N \langle \delta^2 \rangle = N \frac{\langle l^2 \rangle}{3}$$

Now, we make use of the fact that the statistics of run lengths follow Poisson distribution. Thus  $\langle l^2 \rangle = 2 \langle l \rangle^2$ .

$$\langle \vec{x}^2 \rangle = \frac{2n \langle l \rangle^2}{3}$$

As  $\langle l \rangle = v\tau$  ( $\tau$  is the average time taken for each step and  $v$  is a constant) and  $N = \frac{t}{\tau}$ ,

$$\langle \vec{x}^2 \rangle = \frac{2}{3} \frac{t}{\tau} v^2 \tau^2 = \frac{2}{3} v^2 \tau t = \frac{2}{d\alpha} v^2 t = 2D t$$

Here,

$$\alpha = \frac{1}{\tau} \tag{46}$$

$d = 3$  is the dimension of the space and

$$D = \frac{v^2}{\alpha d} \tag{47}$$

is the diffusion constant (Eq.(1) of [4]).

The experimentally measured value of  $\langle l \rangle$  for E.coli is  $20\mu m$ . The diffusivity of live E.coli is hundreds of times larger than the diffusivity of the deflagellated mutant of the same species. The only possible motion they can exhibit then is only Brownian motion which all particles (alive/dead) do.

### Differential equation

Consider a single bacterium confined in the  $x$  axis.  $R(x, t)$  and  $L(x, t)$  are the probabilities of finding the particle at  $(x, t)$  moving rightward or leftward respectively.  $v_L(x)$  and  $v_R(x)$  are the speeds of the bacterium at  $x$  going left or right.  $\alpha_L$  and  $\alpha_R$  are the tumbling rates of the bacterium going left or right respectively. We will write continuity equations for both  $R$  and  $L$  with source and sink terms.

$$\begin{aligned} \dot{R} + (v_R R)' &= -\alpha_R \frac{R}{2} + \alpha_L \frac{L}{2} \\ \dot{L} + (-v_L L)' &= \alpha_R \frac{R}{2} - \alpha_L \frac{L}{2} \end{aligned} \tag{48}$$

The terms on the RHS in the first eq. of (48) can be explained thus. The negative term (sink) indicates that half the particles going towards right when they tumble will go left (remember that tumbling is a random process; there is equal probability to go right/left after a tumble). The positive term (source) indicates that half the particles going left after tumbling will turn right.

Eq. (48) is rearranged to get eq. (3) and (4) of [4].

$$\begin{aligned}\dot{R} &= -(v_R R)' - \alpha_R \frac{R}{2} + \alpha_L \frac{L}{2} \\ \dot{L} &= (v_L L)' + \alpha_R \frac{R}{2} - \alpha_L \frac{L}{2}\end{aligned}\tag{49}$$

The differential equation (when  $\alpha_R = \alpha_L$ ) that the E.Coli bacteria obey can be seen to be the same as Eq.(19) of a persistent random walker.

Though in the simulations of the ensemble of bacteria were not performed in a closed vessel where the system has attained equilibrium, it is instructive to look into the solutions of such cases, to get a feel of how the solutions of the DE's.

These solutions are explained in [9].

$\rho \equiv R + L$ ,  $\sigma \equiv R - L$ ,  $J = v\sigma$ . Consider the case  $v_R = v_L = v(x)$  and  $\alpha_R = \alpha_L = \alpha(x)$ .<sup>4</sup>

The equations in (49) is added and subtracted to get the two equations:

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= -\frac{\partial J}{\partial x} \\ \frac{\partial \sigma}{\partial t} &= -\frac{\partial(v\sigma)}{\partial x} - \alpha\sigma\end{aligned}\tag{50}$$

There is a need to tamper with the second equation in order to get the solution for  $J(x)$ .

$$\begin{aligned}v \frac{\partial \sigma}{\partial t} &= -v \frac{\partial(v\rho)}{\partial x} - \alpha J \\ \frac{\partial}{\partial x} \left( v \frac{\partial \sigma}{\partial t} \right) &= -\frac{\partial}{\partial x} \left( v \frac{\partial(v\rho)}{\partial x} \right) - \alpha \frac{\partial J}{\partial x}\end{aligned}\tag{51}$$

Now, The LHS can be rewritten as:

$$\frac{\partial}{\partial x} \left( v \frac{\partial \sigma}{\partial t} \right) = \frac{\partial}{\partial x} \left( \frac{\partial}{\partial t} (v\sigma) \right) - \frac{\partial}{\partial x} \left( \frac{\partial v}{\partial t} \sigma \right) = \frac{\partial^2 J}{\partial x \partial t}$$

The last term is gone because  $v$  is independent of  $t$ . From the first of (49), we see that  $\frac{\partial^2 J}{\partial x \partial t} = -\frac{\partial^2 \rho}{\partial t^2}$ . So,

$$\frac{\partial}{\partial x} \left( v \frac{\partial \sigma}{\partial t} \right) = -\frac{\partial^2 \rho}{\partial t^2}$$

Substituting in (51), we get

$$\frac{\partial^2 \rho}{\partial t^2} = \frac{\partial}{\partial x} \left( v \frac{\partial(v\rho)}{\partial x} \right) + \alpha \frac{\partial J}{\partial x}\tag{52}$$

This is Eq.(2.4) of [9]. In the steady state,  $\frac{\partial^2 \rho}{\partial t^2} = 0$ . By solving the first order differential equation in  $J(x)$ , we get

$$J(x) = -\frac{v^2(x)}{\alpha(x)} \frac{\partial \rho}{\partial x} - \frac{v(x)\rho(x)}{\alpha(x)} \frac{\partial v}{\partial x}\tag{53}$$

---

<sup>4</sup>Note that the tumbling rates being equal in both directions does not correspond to the observed behaviour of the E.Coli.

- **Case I:**  $v(x) = v$  and  $\alpha(x) = \alpha$  are constant.

Then (127) is,

$$J(x) = -\frac{v^2}{\alpha} \frac{\partial \rho}{\partial x} \quad (54)$$

We see that Fick's law<sup>5</sup> need not be modified. From the first equation of (128), we then get

$$\frac{\partial \rho}{\partial t} = \frac{v^2}{\alpha} \frac{\partial^2 \rho}{\partial x^2} = D \frac{\partial^2 \rho}{\partial x^2} \quad (55)$$

where the diffusion constant  $D = \frac{v^2}{\alpha}$ . Thus we see in this case that the differential equation gives back the unbiased random walk differential equation.

- **Case II:**  $v(x) = v$  is constant,  $\alpha(x)$  is not.

Then,

$$J(x) = -\frac{v^2}{\alpha(x)} \frac{\partial \rho}{\partial x} \quad (56)$$

The *form* of Fick's law is still the same. The equilibrium solution is got by equating  $J(x)$  to 0.

- **Case III:**  $v(x)$ ,  $\alpha(x)$  both are not constant but mean free path  $\frac{v(x)}{\alpha(x)} = c$ . It can be verified that the form of the Fick's law is modified.

$$J(x) = -\frac{\partial(D\rho)}{\partial x} \quad (57)$$

where

$$D = \frac{v^2(x)}{\alpha(x)} \quad (58)$$

The equilibrium solution:

$$J(x) = 0 \Rightarrow D(x)\rho(x) = k$$

Evaluating the constant at the origin, we get  $k = D(0)\rho(0)$ . Thus,

$$\rho = \frac{D(0)}{D} \rho(0) = \frac{v(0)^2}{\alpha(0)} \frac{\alpha(x)}{v^2(x)} \rho(0)$$

where we used  $D(0) = \frac{v(0)^2}{\alpha(0)}$

Knowing that mean free path is a constant:

$$\rho = cv(0) \frac{1}{cv(x)} \rho(0)$$

Thus, the equilibrium distribution of particle density is:

$$\rho(x) = \rho(0) \frac{v(0)}{v(x)} \quad (59)$$

This is the eq.(6) of [4] and Eq.(2.10) of [9].

It can be seen that  $\rho(x)$  is independent of tumbling rate and is more at the places where  $v(x)$  is less. The particles accumulate at such places.

---

<sup>5</sup>It is just a name for the kind of equations of form (54) which relate currents in steady state to the gradient in concentration. Fick in 1855 had postulated that in steady state  $\left(\frac{\partial \rho}{\partial t} = 0\right)$ ,  $J(x) = -D \frac{\partial \rho}{\partial x}$  which is quite intuitive. However, we will be seeing that this form of relation is not true in general.

- **Case IV:**  $\alpha(x) = \alpha$  is constant and  $v(x)$  is a constant.

The Fick's law is different again for this case.

$$J(x) = -\frac{v^2(x)}{\alpha} \frac{\partial \rho}{\partial x} - \frac{v(x)\alpha(x)}{\alpha} \frac{\partial v(x)}{\partial x} = -D \frac{\partial \rho}{\partial x} - \frac{\rho(x)}{2} \frac{\partial D}{\partial x} \quad (60)$$

where

$$D(x) = \frac{v^2(x)}{\alpha} \quad (61)$$

The equilibrium solution for this system is found by similar method.

$$\rho(x) = \sqrt{\frac{D(0)}{D(x)}} \rho(0) = \rho(0) \frac{v(0)}{v(x)} \quad (62)$$

Now, let us generalize the situation to where the tumbling rate and the velocity are direction dependent<sup>6</sup>. In [9], a simple calculation which provides some intuition about the phenomenon is given.

Let the average time for a step to right be  $\tau_R$  and to the left be  $\tau_L$ . These are just the inverse of the tumbling rate  $\alpha_R$  and  $\alpha_L$ . Let's for simplicity take them both to be unequal constants.

Then, average time for  $N$  steps is,

$$T = \frac{N}{2} (\tau_R + \tau_L) \quad (63)$$

as half the times after each tumble the particle goes right and half the other times, left.

Average displacement is

$$\langle x \rangle = \frac{N}{2} (v_R \tau_R - v_L \tau_L) \quad (64)$$

The drift velocity,

$$v_{\text{drift}} = \frac{\langle x \rangle}{T} = \frac{v_R \tau_R - v_L \tau_L}{\tau_R + \tau_L}$$

Now, replacing the  $\tau$  terms by their inverse  $\alpha$ ,

$$v_{\text{drift}} = \frac{v_R \alpha_L - v_L \alpha_R}{\alpha_R + \alpha_L} \quad (65)$$

If say the food is towards the right, then the  $v_{\text{drift}}$  can be towards right by either  $\alpha_R < \alpha_L$  or  $v_R > v_L$ . When both are true, then so much the good for the bacteria.

The  $J(x)$  can be derived in similar method as done previously (as [9] says),

$$J(x) = -\frac{2(v_R + v_L)}{\alpha_L + \alpha_R} \frac{\partial}{\partial x} \left( \frac{v_R v_L \rho}{v_R + v_L} \right) + \left( \frac{\alpha_L v_R - \alpha_R v_L}{\alpha_L + \alpha_R} \right) \quad (66)$$

## 5.2 Simulation

The derivations that were done in the theory part were for the intention of illustration. The simulation was done for an open environment, not a boxed one.

In the simulations done,  $v_R(x) = v_L(x)$  is a constant. The tumbling rate was kept to be 4 when going towards left (lesser food that way) and to be 2 when going right. The step lengths were random and follows a Poisson distribution. To avoid unnaturally long step lengths, their were appropriately truncated.

---

<sup>6</sup>This corresponds to the observed behaviour of E.Coli.

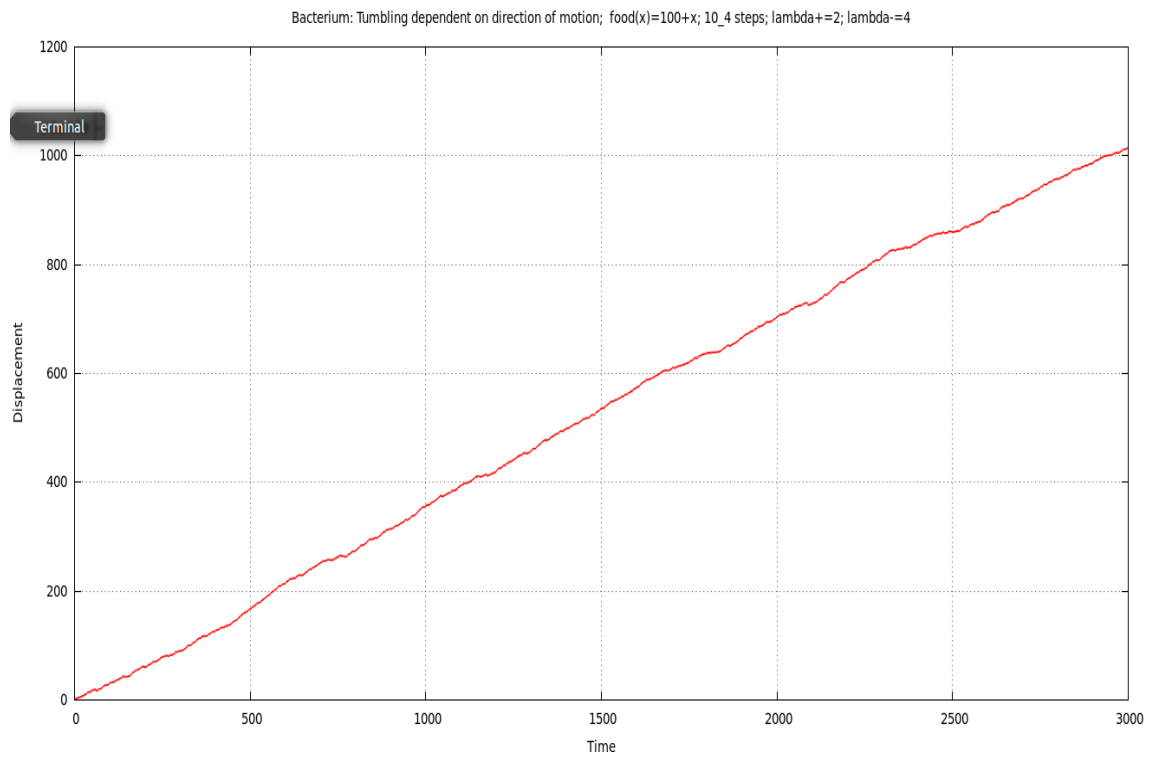


Figure 10: Motion of a single bacterium towards increasing food density

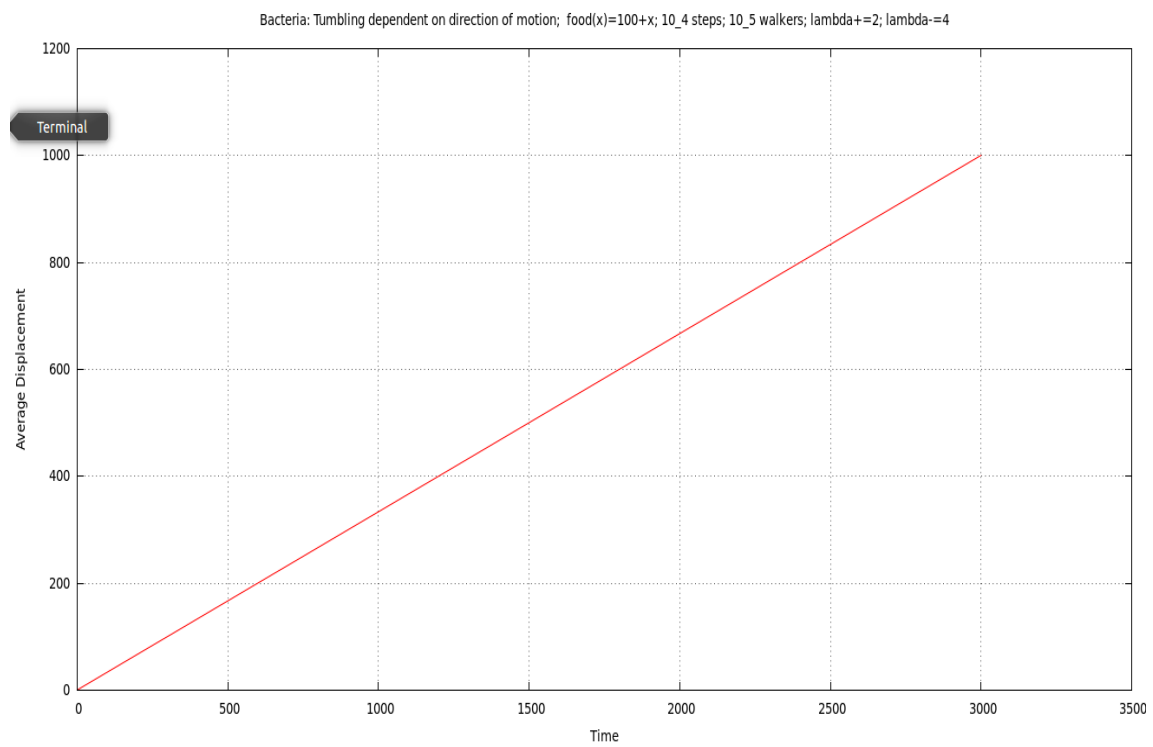


Figure 11: Average displacement of  $10^5$  bacteria over time



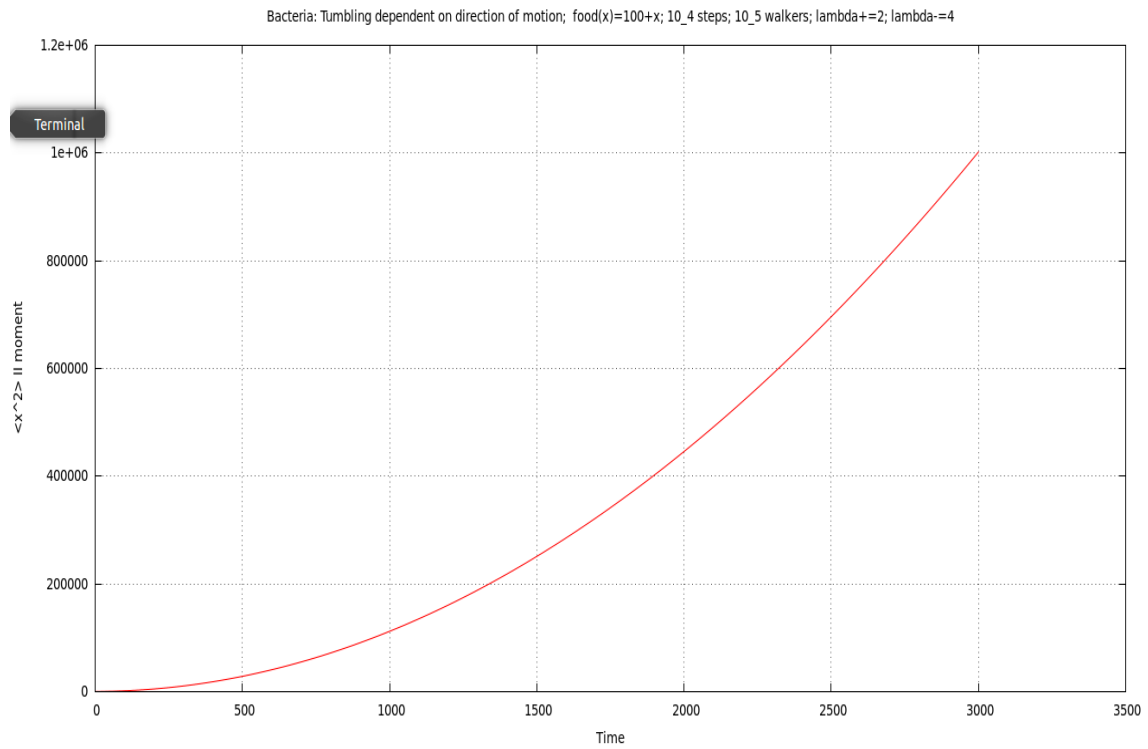


Figure 12: II moment of  $10^5$  bacteria over time

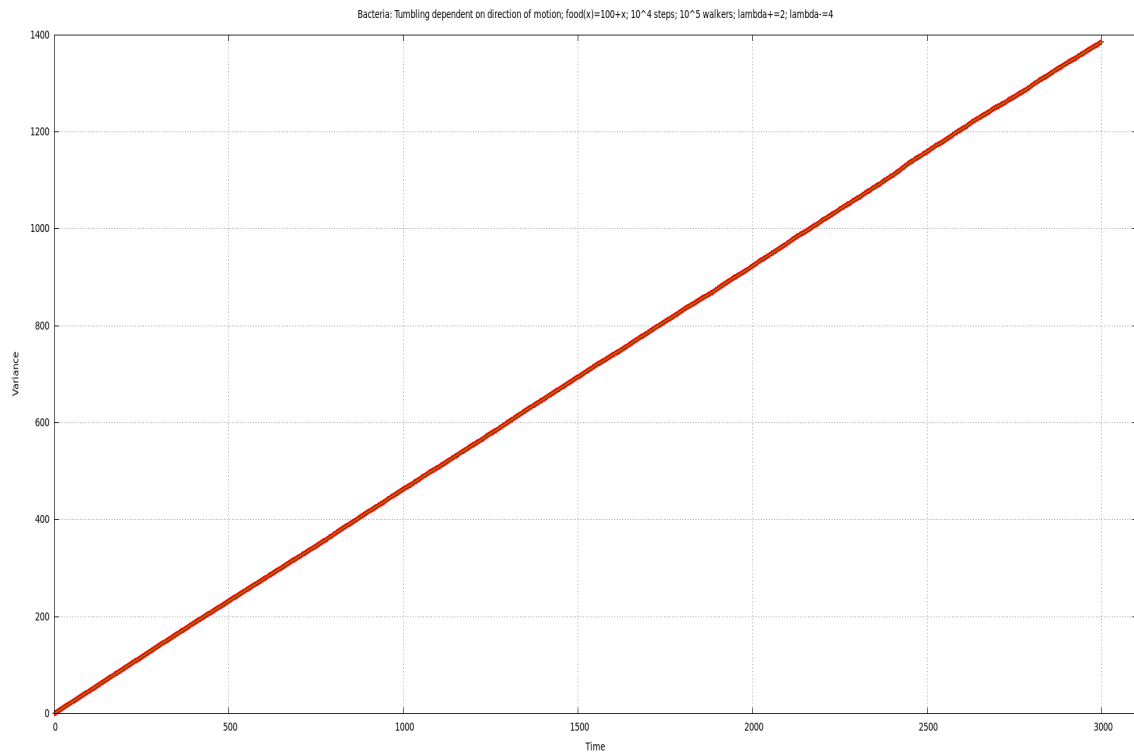


Figure 13: Variance of  $10^5$  bacteria over time

The simulation shows successful chemotaxis when the tumbling rates in different directions are not the same.

## 6 Acknowledgements

I thank Abhishek Dhar for his guidance. I also am grateful to ICTS for housing friendly people and cookie jars.

## 7 References

1. Lecture Notes on Non-equilibrium Statistical Physics - Abhishek Dhar
2. Triple Minima in the Free Energy of the Semiflexible polymers - Abhishek Dhar, Debasish Chaudhuri - Physical Review Letters, Vol.89, No.6
3. Heat and Thermodynamics - Dittman, Zemansky
4. Diffusive transport without detailed balance in motile bacteria: Does micro-biology need statistical physics? - M.E.Cates, arXiv.
5. E. coli in motion - Berg H C - Springer (2003)
6. Kratky, O. and G. Porod. 1949. Rec. Trav. Chim., 68:1106
7. The theory of polymer dynamics - Doi, Edwards - Oxford Science Publications
8. E. Frey et al., cond-mat/9707021.
9. Schnitzer M J 1993 - Theory of continuum random-walks and application to chemotaxis - Phys. Rev. E 48 2553-2568