Random and Annihilating Walks

Byron Adlam

December 2023

Abstract

This report first investigates simple random walks, self-avoiding random walks (SARW) and then a variation on simple random walks; annihilating random walks. In $\mathbb{R}^1$, the distribution of distances is found to tend to a normal distribution. In $\mathbb{R}^2$ it tends to a Rayleigh distribution, where the variance is found to be linear with the step size. For SARWs the variance scales with $N^{1.496}$. For annihilating random walks, the simple case of two walkers is enumerated and the probability that a walker annihilates at step N is found to follow a $\frac{1}{2N}$ distribution. This is found to agree with computational simulations. The model is altered and a new walker is added to the origin every step. The average number of walkers, C_n, per step, N, is found to diverge as $C_n = N^{0.094} \cdot e^{1.43}$.

1 Introduction

Random walks are an interesting stochastic phenomenon that have applications ranging from modelling Brownian motion to polymer conformations. At its simplest, a random walk is a stochastic process in which a 'walker' moves about in a set of predetermined directions. The directions it can move in each step are described by a set of vectors and a corresponding probability distribution. For example, in a simple 1d walk a walker moves either left or right with equal likelihood, 0.5.

To begin with, random walks in $\mathbb{R}^1$ and $\mathbb{R}^2$ are discussed and in the latter, the distribution and variance of distances n steps apart for a square and triangular lattice are generated to gain a better understanding of the simple random walk. Then the self-avoiding random walk (SARW) is examined as an interesting variation. This is a walk that is constrained such that it cannot revisit any points on the lattice. As a result, the variance of distances from the origin is found to scale in a different manner to the simple $\mathbb{R}^2$ random walk. The trapping probability of a SARW is also examined. Finally the annihilating walk is considered, where a base case of is examined first for better understanding. An annihilating walk is characterised by walker elimination at encounters.

2 Random Walk

2.1 1-Dimension

The possible movement vectors for a $\mathbb{R}^1$ random walk are left or right (along a horizontal line).

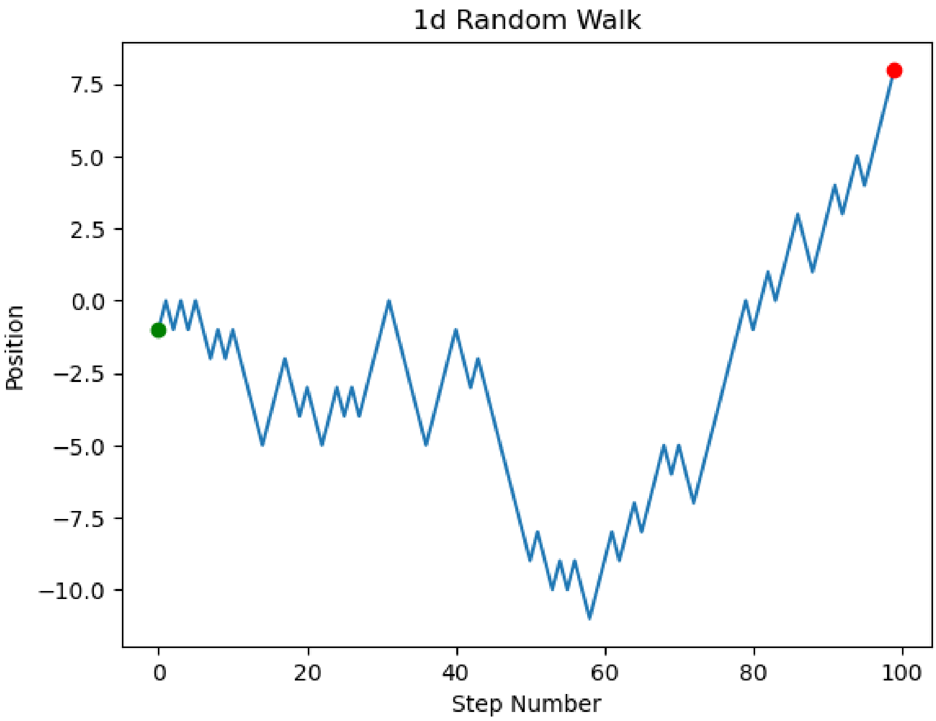

Figure 1: Random Walk in $\mathbb{R}^1$, of length 100

We denote the probability a walker moves left as p, and the probability a walker moves right as $q = 1 - p$. Here we consider a symmetric random walk, and hence $p = q = 0.5$. We can ask ourselves the question, what is the probability, $P(x, N)$, that we find the walker at position x after N moves? Denoting $N = n_1 + n_2$, where n_1 is the number of steps taken to the left and n_2 steps to the right, $P(x, N)$ can be written as;

$$P(x,N) = \frac{1}{2^N} \cdot \binom{N}{n_1} = \frac{1}{2^N} \cdot \frac{N!}{(n_1! n_2!)} \tag{1}$$

Although each random walk is independent of other random walks, if we take a large enough sample size of walks the probability distribution for $P(x, N)$ will tend to a normal distribution, with a mean, $\mu = 0$, and a standard deviation, $\sigma = \sqrt{samplesize}$. This is due to the central limit theorem; a large sample of independently distributed random variables will tend to a normal distribution, even if the individual distributions are non-normal [2].

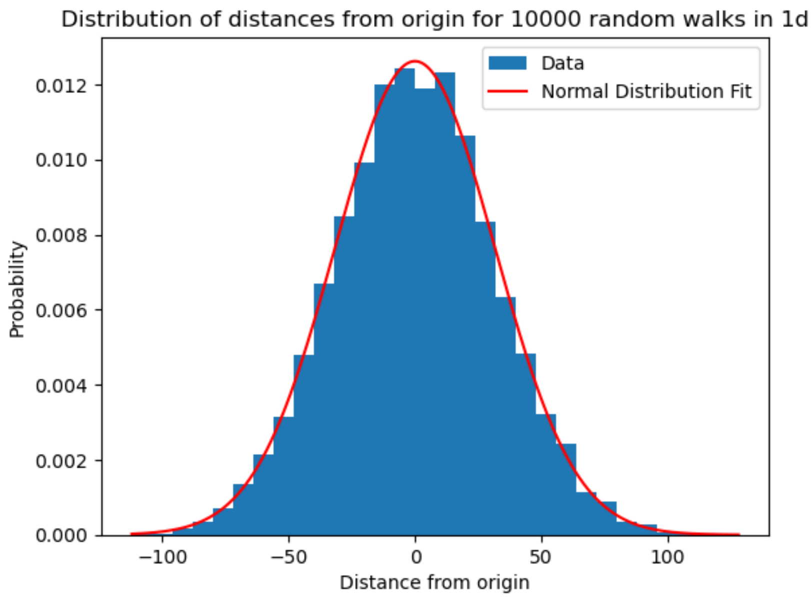

Figure 2: Distribution of distances from the origin for 10000 random walks in $\mathbb{R}^1$

In 2, we can clearly see that for large N, the distribution tends to a normal distribution.

2.2 2-Dimension

2.2.1 Square Lattice

We now examine a random walk on a $\mathbb{R}^2$ square lattice. There are now four movement vectors for a random walker:

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} : \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \tag{2}$$

As a result, the walker has twice the number of directions to choose from when picking its next move.

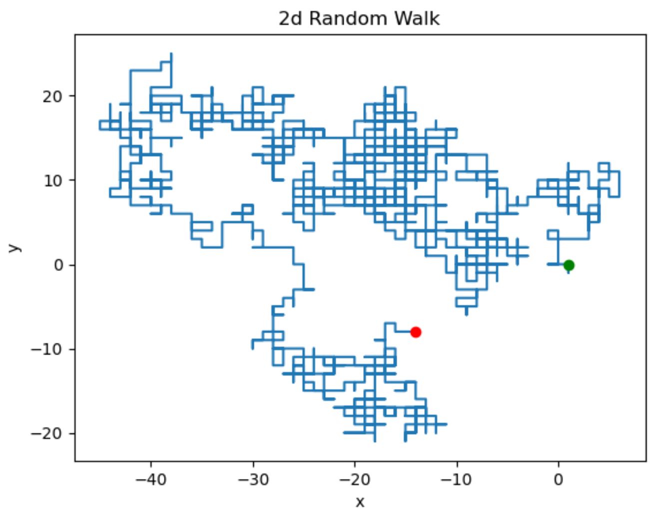

Figure 3: Random walk on $\mathbb{R}^2$ square lattice, of length 2000

To examine the distribution of distances N steps apart, one walk of length 2,000,000 was generated. To fit the distribution, a Chi distribution with two degrees of freedom ($\mathbb{R}^2$) is used [3]. This is more commonly known as a Rayleigh distribution. As the sample size increases, we expect the distribution of points N steps apart to tend towards a Rayleigh distribution:

$$P(x) = \frac{x}{\sigma^2} \exp\left(\frac{-x^2}{2\sigma^2}\right) \tag{3}$$

The variance of a Rayleigh distribution is [3]:

$$Var(X) = \frac{4-\pi}{2}\sigma^2 \tag{4}$$

To find σ, we look at the plot of variance and step size for the distribution of distances.

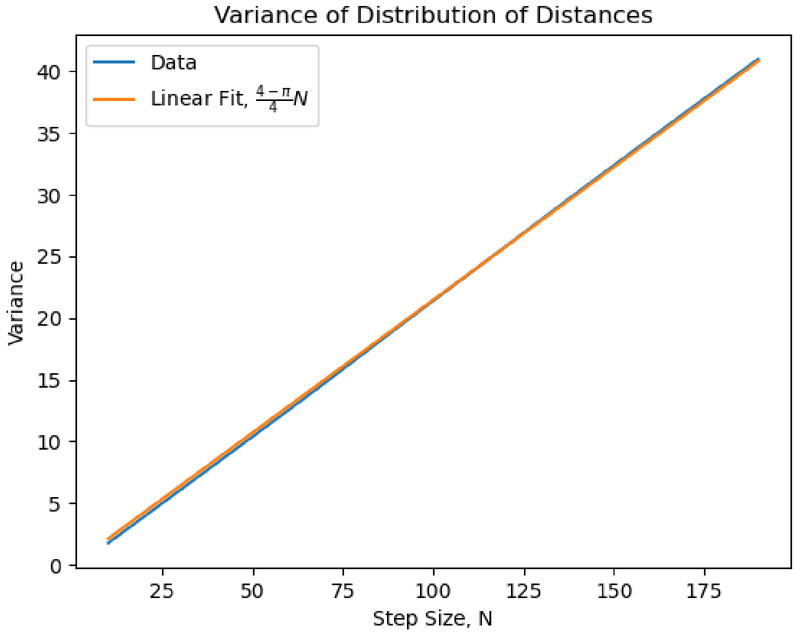

Figure 4: Variance of distribution of distances N steps apart for a random walk on a $\mathbb{R}^2$ square lattice

Here, in 4, it is clear to see the variance varies linearly with step size, following a gradient of $\frac{4-\pi}{4}$. This implies the scale parameter, σ is;

$$\sigma = \sqrt{\frac{N}{2}} \quad \implies Var(N) = \frac{4-\pi}{4}N \tag{5}$$

Seeing how well the Rayleigh variance fits the data, we expect a good fit on the distribution of distances.

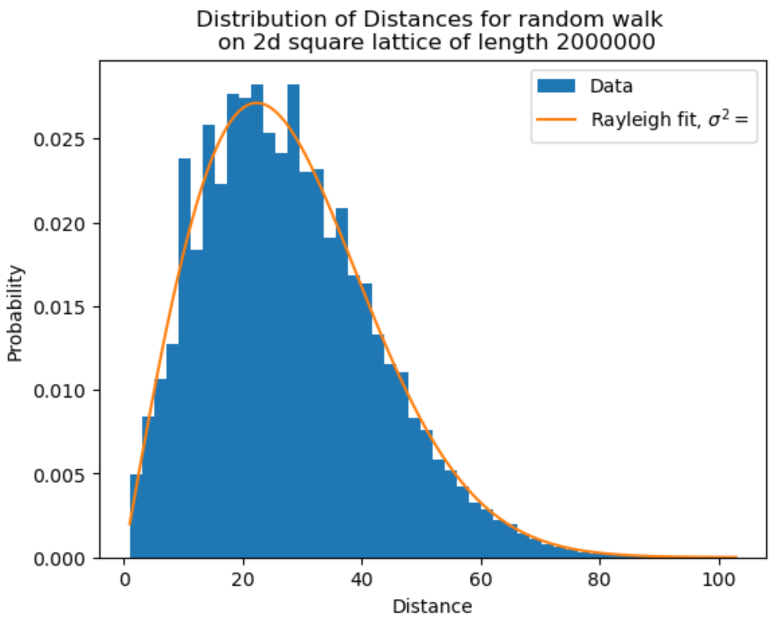

Figure 5: Distribution of distances N steps apart for a random walk on a $\mathbb{R}^2$ square lattice

With a variance of 210.83 in 5, the distribution seems to fit the Rayleigh fit very closely, confirming that the data indeed tends to a Rayleigh distribution.

2.2.2 Triangular Lattice

An interesting variation on the $\mathbb{R}^2$ simple random walk, is performing the walk on a triangular lattice.

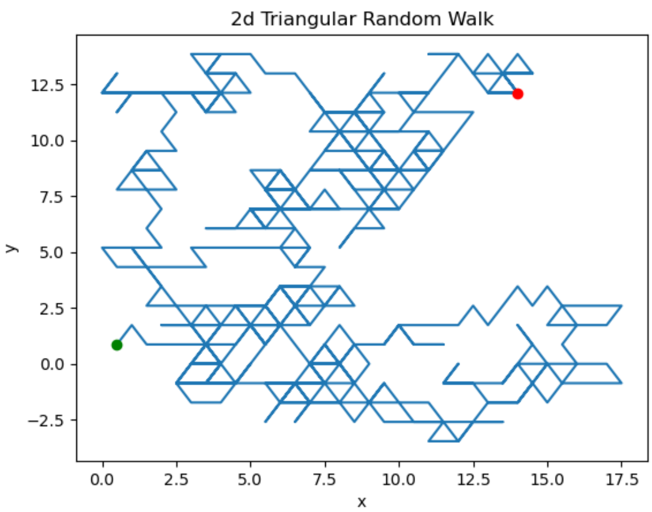

Figure 6: Random walk on a $\mathbb{R}^2$ triangular lattice, of length 500

Here, the possible movement vectors are:

$$\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\end{pmatrix} : \begin{pmatrix}1\\ 0\end{pmatrix}, \begin{pmatrix}-1\\ 0\end{pmatrix}, \begin{pmatrix}\frac{\sqrt{3}}{2}\\ \frac{1}{2}\end{pmatrix}, \begin{pmatrix}-\frac{\sqrt{3}}{2}\\ \frac{1}{2}\end{pmatrix}, \begin{pmatrix}\frac{\sqrt{3}}{2}\\ -\frac{1}{2}\end{pmatrix}, \begin{pmatrix}-\frac{\sqrt{3}}{2}\\ -\frac{1}{2}\end{pmatrix} \tag{6}$$

Interestingly the distribution of distances on a triangular lattice was also found to follow a Rayleigh distribution, indicating the distribution is independent of the lattice structure. The Variance was calculated and was also found to vary linearly with N.

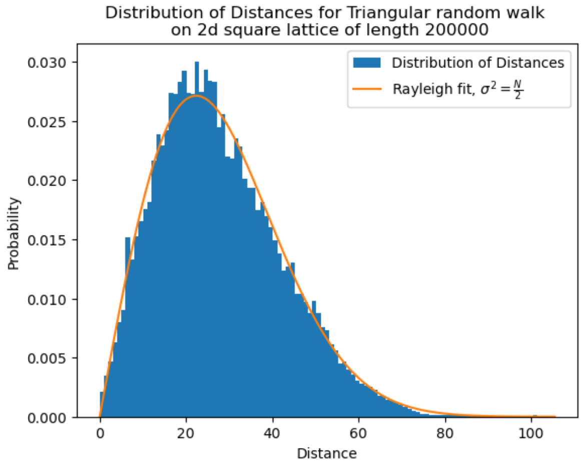

Figure 7: Distribution of distances N points apart for a random walk on a $\mathbb{R}^2$ triangular lattice

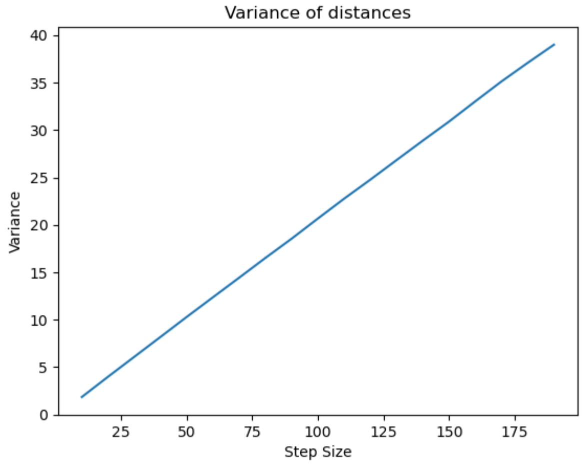

Figure 8: Variance of distribution of distances for a random walk on a $\mathbb{R}^2$ triangular lattice

A value of 211.5 for the variance was calculated from the distribution.

3 Self-Avoiding Random Walk (SARW)

A self-avoiding random walk, SARW, is a random walk which does not revisit any point on the lattice that it has already visited. It is defined as a trapped walk when it can no longer move in a direction without revisiting a point.

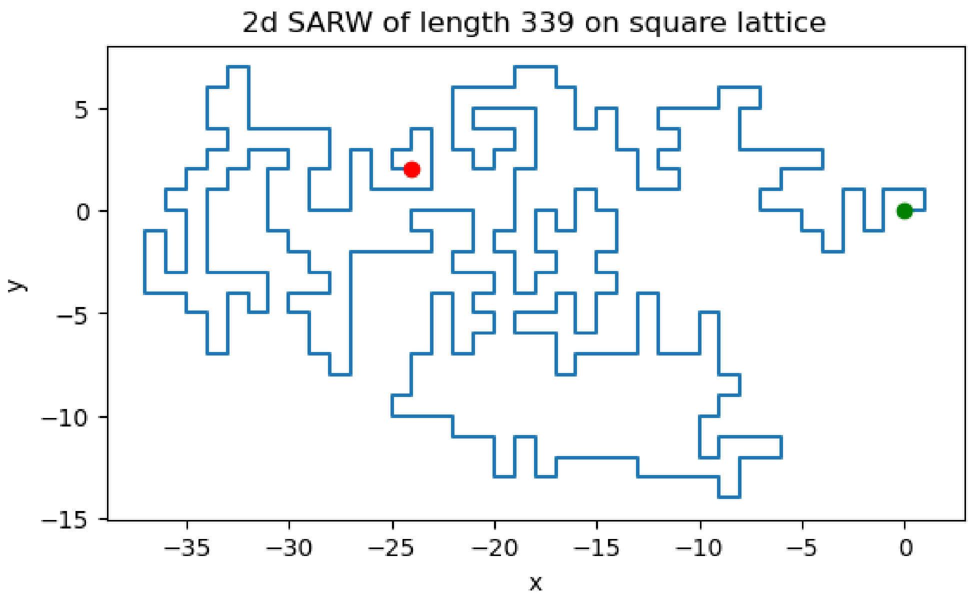

Figure 9: Self-Avoiding Random Walk of length 339

3.1 Algorithm

The algorithm for generating a SARW is different to that of a simple random walk. Due to the SARW's revisiting constraint, the algorithm is more complex. The walker starts at the origin of the sample space, in the $\mathbb{R}^2$ case $\binom{\mathbf{x}}{\mathbf{y}} = \binom{0}{0}$, and for each step it takes, it chooses a movement vector from eq. 2. The chosen vector is then compared to all previous positions of the walker. If the chosen vector results in a position already visited, it is removed from the available movement vectors and the walker chooses another. This is repeated until either an available position is found or the movement vectors are exhausted. If the former occurs, the movement vectors are reset, the position is updated and the step is incremented. If the latter occurs, the walker is described as trapped and the algorithm ends.

3.2 Variance

For random walks the variance has been found to be linearly proportional to the number of steps. However, for a SARW the variance is postulated to scale with N as [1]:

$$\sigma^2 = DN^{2\nu} \tag{7}$$

Where D is a constant, and ν is a critical exponent. In $\mathbb{R}^2$, ν is accepted to be $\frac{3}{4}$. To generate a distribution of distances for random walks, one long walk of 2,000,000 steps could be created and split into sections N steps apart. This is because the fractal nature of a random walk means the walk between X_1 and X_2 displays the same statistical properties as between points X_2 and X_3. However, for a SARW, after it has travelled from X_1 to X_2 (assuming it does not trap), the walk from X_2 and X_3 now has a reduced sample space to travel in, and thus it has different statistical properties as it cannot revisit points it has already traversed. Instead, to generate a distribution of distances, 50,000 SARWs were generated and the variances of distribution of distances of walks per step N were found.

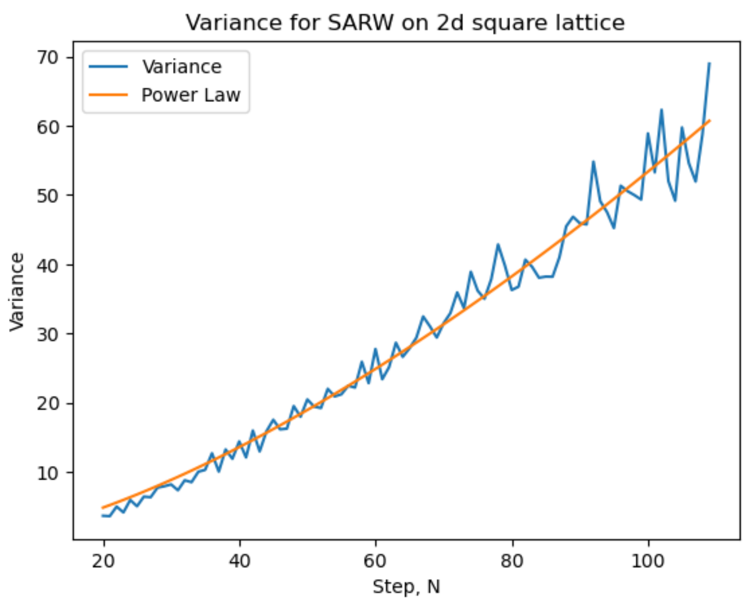

Figure 10: Variance for distribution of distances per step, N for a SARW in $\mathbb{R}^2$

The equation for the power law fit in fig.10 is:

$$Var(N) = 0.0544N^{1.496} \tag{8}$$

The fit was calculated using scipy.curve_fit. The power law equation seems to fit the data very well and gives a value of $\nu = 0.748$ for the critical exponent, which is very close to the accepted value of $\frac{3}{4}$ [4].

3.3 Trapping Probability

The trapping nature of SARWs is investigated further, specifically the probability that a SARW will trap at step $N > 6$. For this, the 50,000 random walks that were generated previously, for the variance, are used. Then the number of walkers that trap at N steps is counted, and thus the probability determined. $N > 6$ is chosen as the walk must take a minimum of 6 steps before it can trap.

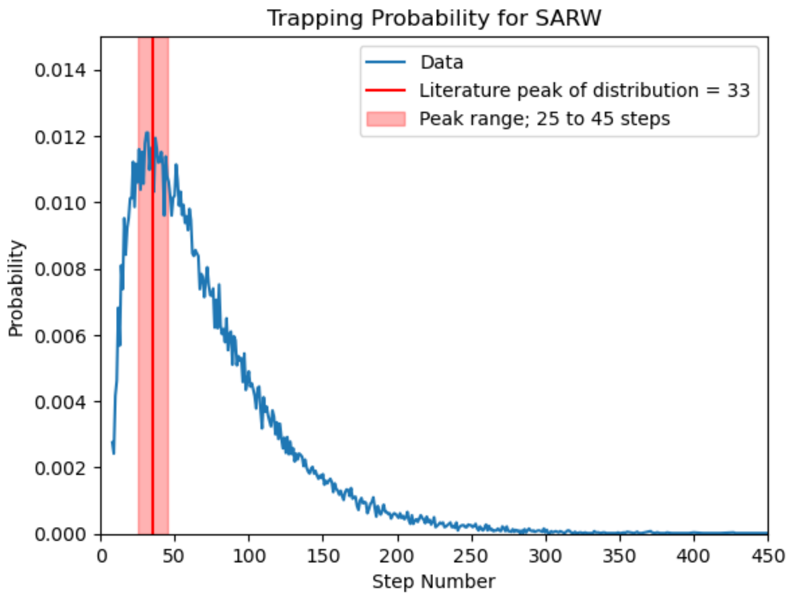

Figure 11: Trapping probability of a SARW per step in $\mathbb{R}^2$

The peak of the probability distribution is represented by the red-shaded region of fig.11. The maximum value obtained by the data is at 31 steps. This value is close to the literature value of 33 [1]. This indicates that the most likely number of steps for a SARW to travel before trapping is between 25 and 45 steps. After this the probability decreases rapidly and at 350 steps it is almost negligible. The mean number of steps taken before trapping was also calculated, and a value of 71.5 steps was returned. This value is also very close to the literature value of 70.7 [1].

4 Annihilating Walks

Annihilating walkers are random walkers who are removed from the sample space when they occupy the same position at the same. The model that is investigated is described as follows. A walker begins at the origin of the sample space. Every time it takes a step in the sample space as described by a simple random walk, a new walker is placed at the origin. If two of these walkers occupy the same position at the same time, they 'annihilate' each other and are removed. To examine this model, a base case is investigated first.

4.1 Base Case

We begin with two walkers at the origin of a $\mathbb{R}^1$ sample space, and no new walkers being added at each step. If the two walkers annihilate, the program is terminated. This sample space for two annihilating walkers is simple enough to be enumerated.
If we expand the possible positions a walker can travel to in $\mathbb{R}^1$, we get a pascals triangle.

Step Number, N	Walker Positions
N=0	●
N=1	● ●
N=2	● ● ●
N=3	● ● ● ●
N=4	● ● ● ● ●
N=5	● ● ● ● ● ●
N=6	● ● ● ● ● ● ●

Table 1: Pascal's triangle representing possible starting positions of random walkers per step N

We want to examine the probability that two walkers annihilate at a step, N, given that they haven't yet annihilated, which we denote $P(Annihilate, N)$. To calculate this, we will need the number of annihilations per step and the total positional combinations for two walkers at that same step:

$$P(Annihilate, N) = \frac{\text{Number of Annihilations per Step, N}}{\text{Total Combinations of Positions for Two Walkers}}$$

Let us look at the first two steps.

At step N=1 we have two possible starting positions for the walkers. Denoting each starting position by a and b respectively, and a move to the left or right as subscript L and R, we can have the following moves:

$$\begin{pmatrix} \mathbf{X}_{W_1} \\ \mathbf{X}_{W_2} \end{pmatrix} : \begin{pmatrix} a_L \\ b_L \end{pmatrix}, \begin{pmatrix} a_L \\ b_R \end{pmatrix}, \begin{pmatrix} a_R \\ b_L \end{pmatrix}, \begin{pmatrix} a_R \\ b_R \end{pmatrix} \tag{9}$$

Here only $\left(\begin{smallmatrix} a_R \\ b_L \end{smallmatrix}\right)$ results in an annihilation, assuming we evaluate everything from left to right; this gives $P(Annihilate, 1) = \frac{1}{4} = 0.25$. Now $N = 2$ and there are three possible starting positions for the random walkers (see Table 1). Denoting the new starting positions by a, b and c respectively gives the following moves:

$$\begin{aligned} \begin{pmatrix} \mathbf{X}_{W_1} \\ \mathbf{X}_{W_2} \end{pmatrix} : \begin{pmatrix} a_L \\ b_L \end{pmatrix}, \begin{pmatrix} a_L \\ b_R \end{pmatrix}, \begin{pmatrix} a_L \\ c_L \end{pmatrix}, \begin{pmatrix} a_L \\ c_R \end{pmatrix}, \begin{pmatrix} a_R \\ b_L \end{pmatrix}, \begin{pmatrix} a_R \\ b_R \end{pmatrix}, \begin{pmatrix} a_R \\ c_L \end{pmatrix}, \begin{pmatrix} a_R \\ c_R \end{pmatrix}, \\ \begin{pmatrix} b_L \\ c_L \end{pmatrix}, \begin{pmatrix} b_L \\ c_R \end{pmatrix}, \begin{pmatrix} b_R \\ c_L \end{pmatrix}, \begin{pmatrix} b_R \\ c_R \end{pmatrix} \end{aligned} \tag{10}$$

Here $\binom{a_R}{b_L}$ and $\binom{b_R}{c_L}$ annihilate, giving $P(Annihilate, 2) = 2/12 = 0.1\dot{6}$. It is clear to see number of annihilations per step (row in pascal's triangle) is;

$$\text{Number of Annihilations per Step} = N - 1,$$

as an annihilation requires walkers to come from adjacent positions. At each step, the possible positional combinations for two walkers is;

$$\text{Number of positional combinations} = 2 \cdot 2 \cdot \binom{N}{2} = 2N(N-1)$$

As we have two walkers who can each go in two different directions (left or right) with the same probability, and there are $\binom{N}{2}$ combinations of two walkers from N, where the step number represents the number of positions a walker can travel from. Therefore, the probability two walkers annihilate at step N given that they haven't yet annihilated is;

$$P(Annihilate, N) = \frac{1}{2N}$$

n.b. we have a special case at N=0, as at the start we allow the two walkers to exist together at the same point. Here the number of annihilations is actually 2, and the number of positional combinations is 4. Hence $P(Annihilate, 0) = 0.5$.

As a result, calculating the probability for two walkers to annihilate per step given that they haven't yet annihilated, should yield the following plot (fig. 12).

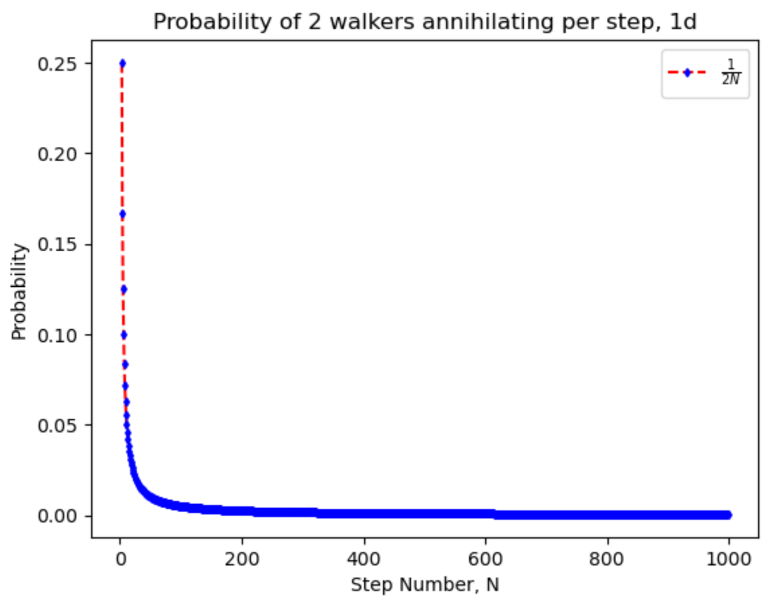

Figure 12: Probability of two walkers annihilating per step in $\mathbb{R}^1$, enumerated analytically

Now by running 10,000 simulations of this base case, we can compare fig. 12 with the computational data (fig. 13). Here we see the data clearly displays a $\frac{1}{2N}$ fit, with some spread in the data around the lower values seen in the log-log plot (fig. 14) as can be expected when fitting simulated data. This confirms the model that $P(Annihilate) = \frac{1}{2N}$.

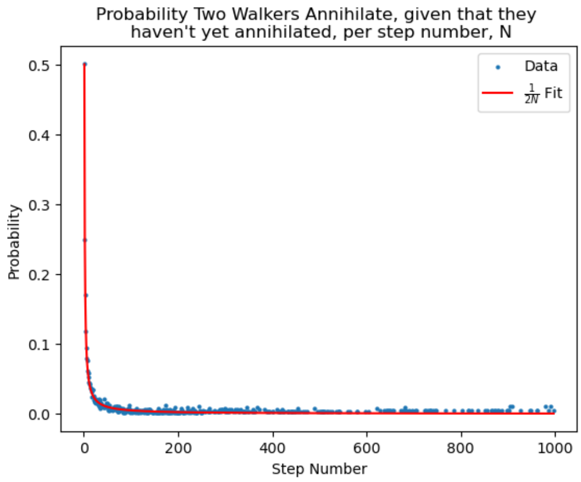

Figure 13: Probability two $\mathbb{R}^1$ walkers annihilate per step number N, given that they haven't yet annihilated

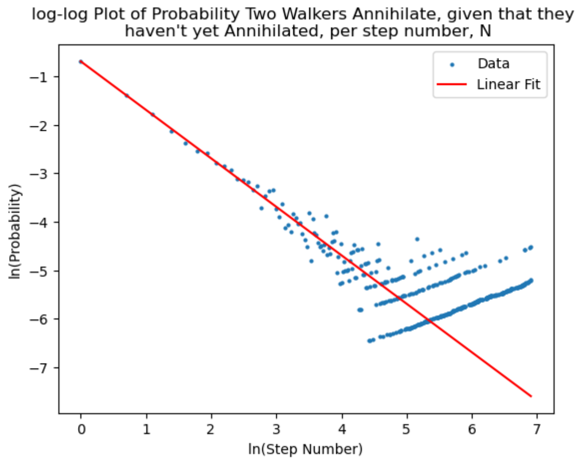

Figure 14: log-log plot of the probability of two $\mathbb{R}^1$ walkers annihilating per step N, given that they haven't yet annihilated

The distribution of the number of steps two walkers take also provides some insight.

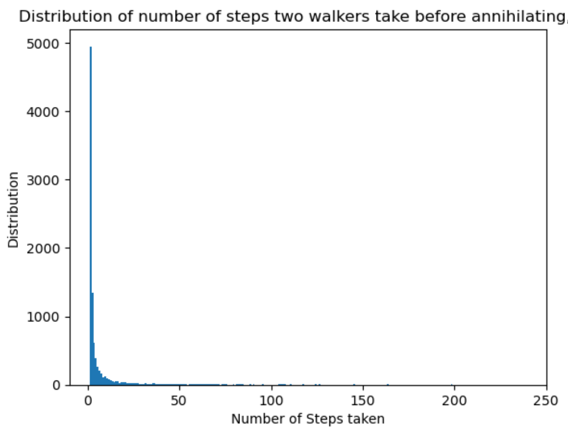

Figure 15: Distribution of number of steps two walkers take before annihilating in $\mathbb{R}^1$

Here, in fig. 15, a small mean of 18.4 steps and a large variance of 6122.2 steps indicates that although there is a wide spread of durations of walks, the majority are found close to the origin. This is further supported by the shape of the distribution, specifically the rate at which the distribution per step decreases at small steps. n.b., the x-axis (number of steps taken) in fig. 15 is truncated to make the distribution clearer. The number of steps taken has values at 1000, however the corresponding distribution is so small it cannot be seen. Thus to better see the distribution shape in the fist few steps the axis was truncated.

Having now examined the simple base case for two annihilating walkers, we can move onto the main model.

4.2 Main Model

As described in section 4, we start with one walker at the origin, and a random walker is added to the origin every step. We examine the average number of walkers, C_n, per step,

N, in $\mathbb{R}^1$ to see how it behaves; whether it converges or diverges.

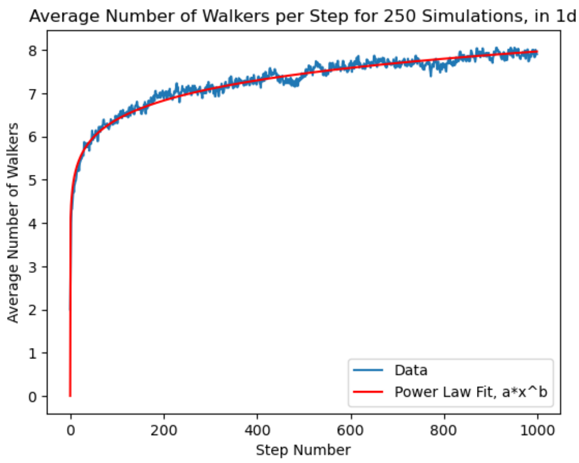

Figure 16: Average number of walkers per step for 250 simulations, in $\mathbb{R}^1$

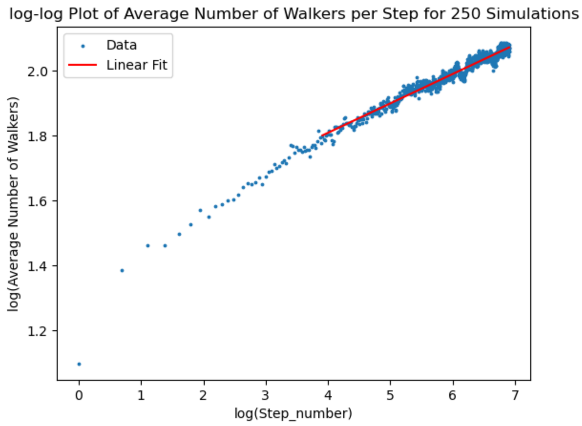

Figure 17: log-log plot of average number of walkers per step for 250 simulations, in $\mathbb{R}^1$

We see in fig. 16, the average number of walkers per step follows a power law fit of

$$C_n = 4.1177N^{0.0955} \tag{11}$$

very nicely. The power law fit indicates the number of walkers per step diverges very slowly; as N tends to infinity, so does C_n. This tells us that the existing walkers do not annihilate faster than the one walker being added in per step. At a small step number, $N < 150$, C_n increases very quickly, as the density of walkers is too small for many annihilations to occur. As N increases thereafter, the total density of walkers increases and it becomes more likely for two walkers to annihilate. In fig. 17, the log-log plot gives a linear fit of

$$\ln C_n = 0.0905 \ln N + 1.4471 \tag{12}$$

for data points of $N > 50$. The difference in the power from the power law fit and the gradient of the linear fit from the log-log plot is due to the latter being truncated by the first 50 steps. This is to ensure that only the linear parts of the plot are fitted. This can be seen later in the correction plot in fig. 19. The gradient of the linear fit from the log-log graph describes how the minimum area of the sample space would have to vary to maintain the same density of walkers. The value of 0.0905 again confirms that the walkers annihilate each other at a sublinear rate compared to N, hence the divergence. We can also infer that, compared to the walker that has travelled the furthest at step N, the other walkers are annihilating each other very quickly.

The sublinear growth of average number of walkers is further supported by a plot of the total density of walkers per step.

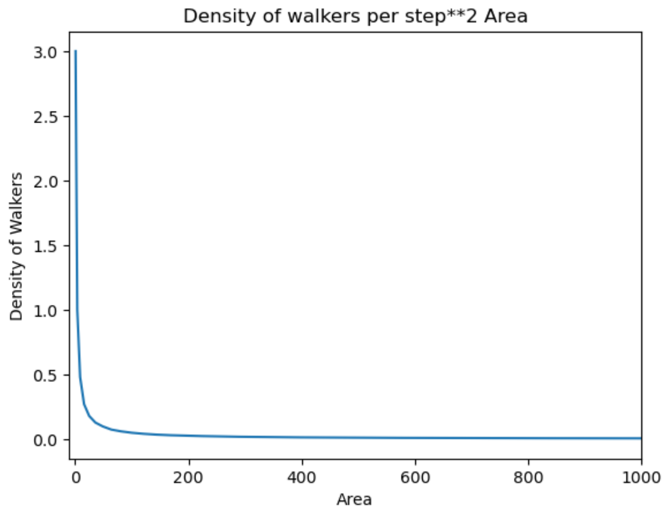

Figure 18: Density of walkers per step^2 area, $\mathbb{R}^1$

Here in fig. 18, the density of walkers drops sharply straight after the origin, supporting the fact the average number of walkers varies much slower than N. The x-axis in fig. 18 has been truncated from 1,000,000 to 1,000 to better show the change in density. After 1,000, the density is negligible.

Dividing C_n by our power law fit, equation 12, gives the following correction plot:

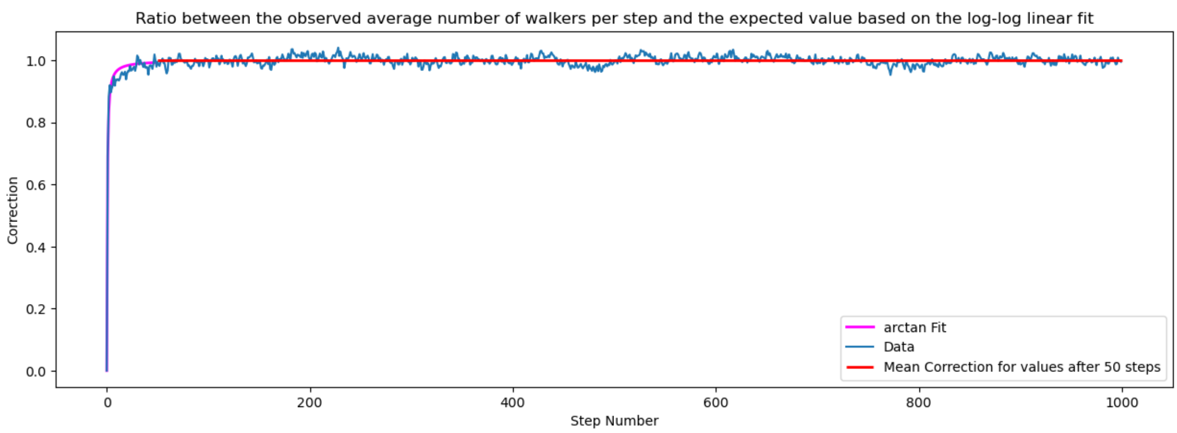

Figure 19: Correction plot for average number of walkers per step in $\mathbb{R}^1$

This correction term clearly tends to a value of 1 as shown by a mean of 1.0001 for $N > 100$. This tells us that:

$$\lim_{N\to\infty} \frac{C_n}{N^{0.0905} \cdot e^{(1.4471)}} = 1 \tag{13}$$

I.e. the function for C_n asymptotes to:

$$\implies C_n = \frac{C_n}{N^{0.0905} \cdot e^{(1.4471)}} \tag{14}$$

This further confirms that the power law fit, for the average number of walkers per N steps, fits very well. We can also fit a function to the first part of the correction plot (arctan fit in fig. 19), for $N \leq 100$:

$$f(N) = \frac{2}{\pi \arctan(1.889N)} \tag{15}$$

This could represent further higher order correction functions for the above equation, however, this fit is not founded on any theoretical model.

4.3 Scaling Dimensions

The model is further examined in $\mathbb{R}^2$, $\mathbb{R}^3$ and $\mathbb{R}^4$. Comparing the gradients of the linear fits for the log-log plots for average number of walkers per step in the first four dimensions:

$$\text{log-log gradient} = \begin{cases} 0.0905; & \mathbb{R}^1 \\ 0.2682; & \mathbb{R}^2 \\ 0.5829; & \mathbb{R}^3 \\ 0.8586; & \mathbb{R}^4 \end{cases} \tag{16}$$

As the dimension increases the possible directions a walker can move in increases, hence making it less likely for two walkers to annihilate. As a result, one would expect there to be more walkers in the same sized area for higher dimensions. Thus, to maintain a minimum area of constant density (what the gradient describes as discussed in section 4.2), the gradient must increase each time the dimensionality is increased. This is seen clearly in equation 16; the gradient increases each time $\mathbb{R}$ increases. The important thing to compare here is the gradient with the general case of no annihilating walkers. As $\mathbb{R}$ increases, we actually hypothesise that the gradient will tend to 1. Since it is much less likely for two walkers to annihilate at higher dimensions, the number of walkers per step will tend towards a model where no annihilation occurs; i.e. every step a new walker is added and none are removed, so the number of walkers per step varies linearly as N. In equation 16, the gradient clearly increases each time. Although the change in gradient from $\mathbb{R}^2$ to $\mathbb{R}^3$, 0.3147, is larger than from $\mathbb{R}^1$ to $\mathbb{R}^2$, 0.1777, the change from $\mathbb{R}^3$ to $\mathbb{R}^4$, 0.2757, is smaller. This could support the fact that the gradients will tend to 1, however, more simulations at higher $\mathbb{R}$ would be needed to verify this. The investigation was stopped at $\mathbb{R}^4$, as the algorithm is too expensive to run past this point. Here, the quality of the random number generator being used affects the statistical properties of walks.

5 Conclusion

The distribution of distances N steps apart was found to tend to a normal distribution in $\mathbb{R}^1$, and display a Rayleigh distribution in $\mathbb{R}^2$ for both a square and triangular lattice. The variance was found to scale linearly with N for both the square lattice and the triangular lattice. The variance of a self-avoiding random walk (SARW) was found to scale as $N^{1.496}$ which is very close to the postulated value of 1.5 [4]. The trapping probability per step, N, was also calculated, where the peak of the probability distribution, 31 steps, was found to agree with the literature value of 33 [1]. Finally a model of annihilating walkers was investigated. First the probability of two walkers annihilating per step given that they haven't yet annihilated, was enumerated and then confirmed computationally to be $P(Annihilate, N) = \frac{1}{2N}$. Then the average number of walkers, C_n, per step, N

was found to asymptote to $C_n = N^{0.0905} \cdot e^{1.4471}$, where the value 0.0905 indicates the how the minimum area surrounding the walkers would have to vary in order to maintain a constant density. Finally the model was scaled up in dimension and found that this value for the gradient, indicating the change in size of minimum area, increases. It is hypothesised that this value will tend to 1; as the $\mathbb{R}$ increases, annihilations will become less likely, and the average number of walkers per step will tend to a model with no annihilations. However, further computations would be required to confirm this.

References

[1] S. Hemmer and P. C. Hemmer. "An average self-avoiding random walk on the square lattice lasts 71 steps". In: *The Journal of Chemical Physics* 81.1 (July 1984), pp. 584–585. ISSN: 0021-9606. DOI: 10.1063/1.447349. eprint: https://pubs.aip.org/aip/jcp/article-pdf/81/1/584/9721538/584_1_online.pdf. URL: https://doi.org/10.1063/1.447349.

[2] Mitch Hill. "APPROXIMATING THE RANDOM WALK USING THE CENTRAL LIMIT THEOREM". In: (Aug. 2011).

[3] Chris H. Rycroft. "Lecture 1: Introduction to Random Walks and Diffusion". In: (Feb. 2005).

[4] Gordon Slade. "The self-avoiding walk: A brief survey". In: 2011. URL: https://api.semanticscholar.org/CorpusID:14773001.