MASS Colloquium

Random Walks: simple and self-avoiding

Notes of the talk given by Gregory F. Lawler
By Hongxu Wei, John Hirdt and Xiaolan Yuan

ABSTRACT: Many phenomena are modeled by walkers that wander randomly. The case of complete randomness is well understood – I will survey some of the key facts including the idea that the set of points visited by a random walker in any dimension (greater than one) is two. I will then discuss a much harder problem – what happens when you do not allow the walker to return to points? Many of the interesting questions about this "self-avoiding walk" are still open mathematical problems.

1 One-dimensional simple random walk

A random walker moves one step to the right or one step to the left with equal probability (1/2) at each time. The position of the walker after n steps S_n is given by the equations below, where X_i and X_j are independent when $i \neq j$.

$$S_0 = 0$$

$$S_n = X_1 + X_2 + \cdots + X_n$$
$$X_j = \begin{cases} +1 & \text{prob } \frac{1}{2} \\ -1 & \text{prob } \frac{1}{2} \end{cases}$$

2 Questions

There are many questions that can be asked in regard to random walks. Some of these have answers that are easy to find, while others are much more difficult. Some of the questions are

- How far does the walker go in n steps?
- Does the walker keep returning to the starting point?

Let $\mathbb{E}(\cdot)$ denote expectation (mean, average value). Expectation is a kind of integral. By the property of symmetry we can see that $\mathbb{E}(S_n) = 0$. We can then ask what about $\mathbb{E}(|S_n|)$? It turns out that it is easier to compute $\mathbb{E}(S_n^2)$. The calculation uses an important property of expectation, linearity. If a, b are numbers, then

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

$$\mathbb{E}[S_n^2] = \mathbb{E}[(X_1 + \cdots + X_n)^2]$$
$$= \mathbb{E}[X_1^2 + \cdots + X_n^2 + a\ lot\ of\ terms\ like\ X_j X_k]$$
$$= \mathbb{E}[X_1^2 + \cdots + \mathbb{E}[X_n^2] + a\ lot\ of\ terms\ like\ \mathbb{E}[X_j X_k]$$
$$= 1 + \cdots + 1 + (a\ lot\ of\ 0's) = n$$

The equations below help in the simplification of the equations above.

$$\mathbb{E}[X_j^2] = \frac{1}{2} \cdot 1^2 + \frac{1}{2} \cdot (-1)^2 = 1$$

$$\mathbb{E}[X_j X_k] = \frac{1}{4} \cdot 1 \cdot 1 + \frac{1}{4} \cdot 1 \cdot (-1) + \frac{1}{4} \cdot (-1) \cdot 1 + \frac{1}{4} \cdot (-1) \cdot (-1) = 0$$

From the equations above, it follows that the walker tends to be approximately distanced $n^{\frac{1}{2}}$ from its starting point. This " $\frac{1}{2}$-law " is fundamental in probability. It states that for n independent trials, the error coming from randomness of order $n^{\frac{1}{2}}$. It is easiest to consider S_n^2, and we know that independence is analogous to orthogonality and one can consider the computation as a generalized form of the Pythagorean theorem

$$a^2 + b^2 = c^2.$$

If we were to look at this in higher analysis, the analogue is L^2, which is a Hilbert space, or an inner product space. The probability distribution of $\frac{S_n}{\sqrt{n}}$ above approaches the normal distribution, more precisely we get

$$\lim \mathbb{P}(S_n \leq a\sqrt{n}) = \int \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

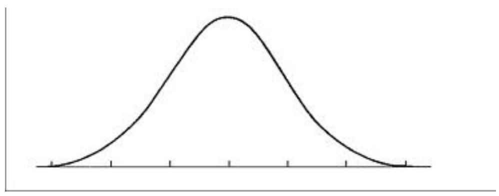


Figure 1: A simple bell curve to demonstrate the normal distribution

This is a case of the central limit theorem – the average of independent samples approaches a normal distribution.

We pose the question, what is the probability of being back at the origin after $2n$ steps? We know that there are $\binom{2n}{n}$ choices of which n steps will be $+1$ and the rest will be -1.

$$\mathbb{P}(S_{2n} = 0) = \binom{2n}{n} (\frac{1}{2})^{2n} = \frac{(2n)!}{n!n!} 4^{-n}.$$

This formula is not so easy to use for direct computation. To find out what it looks like for large n, we use Stirling's formula

$$n! \sim \sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}}.$$

Substituting this in the previous formula gives

$$\mathbb{P}(S_{2n} = 0) \sim \frac{1}{\sqrt{\pi n}}.$$

The power law $n^{-1/2}$ can be seen as follows. The walker tends to be within distance $n^{1/2}$ from the origin and there are $O(n^{1/2})$ such points. So, ignoring the constants, the probability of being at a particular one looks like $n^{-1/2}$.

We can also ask the question: Does the walker keep returning to the origin? Let V be the total number of visits to 0. Then $V = Y_0 + Y_2 + Y_4 + \cdots$, where

$$Y_{2n} = \begin{cases} 1 & S_{2n} = 0 \\ 0 & S_{2n} \neq 0 \end{cases}$$

$$\mathbb{E}[Y_{2n}] = \mathbb{P}(S_{2n} = 0) \sim cn^{-1/2}$$

$$\mathbb{E}[V] = \mathbb{E}[Y_0] + \mathbb{E}[Y_4] + \cdots = \infty \quad \text{since}$$

$$\sum_{n=1}^{\infty} n^{-1/2} = \infty \quad \text{as} \quad \sum_{n=1}^{\infty} n^{-q} \begin{cases} = \infty & q \leq 1 \\ < \infty & q > 1 \end{cases}$$

And thus the random walker returns to the starting point infinitely often.

3 Many dimensions

What will happen in higher dimension?

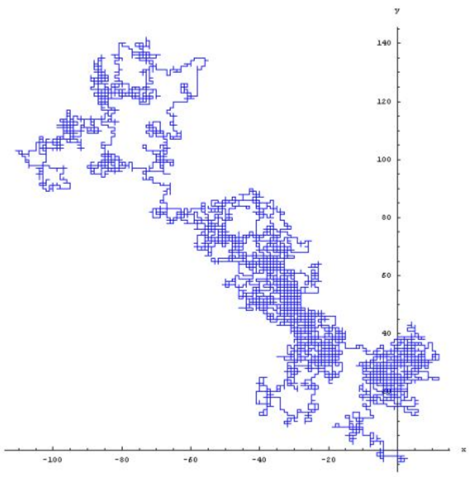


Figure 2: random walk on 2 dimensional integer grid

Generally, we can talk about random walk in d dimensions. In 1 dimension, we have taken probability of $50\% - 50\%$ for the two directions. And in d dimensional integer grid, we have $2d$ directions. Thus in each time step, the walker is equally likely to choose any of $2d$ nearest neighbors. And in these rules, the "square root" law still applies – the typical distance from the starting point after n steps looks like $n^{\frac{1}{2}}$. In d dimensions, the number of points within distance $\sqrt{n}$ of the origin is about $cn^{\frac{d}{2}}$. From this, we can see that the probability being back to the origin at $2n$ steps in d dimensions is $P\{S_{2n} = 0\} \sim c'n^{-\frac{d}{2}}$, and the sum is

$$\sum_{n=0}^{\infty} P\{S_{2n} = 0\} \;=\; \begin{cases} = \infty, & d \leq 2 \\ < \infty, & d > 2 \end{cases}$$

Theorem 1 (Polya)**.** *If $d = 1, 2$, the random walker returns to the origin infinitely often. If $d \geq 3$, the random walker returns only a finite number of times.*

One interesting example of the simple random walk is the drunkard's walk. Imagining you have a drunkard, and the question is, if the drunkard walks around, will he keep returning to the origin point? The answer is 'Yes' in 2 dimensions. But if you give the drunkard wings, actually you consider it in 3 dimensions, then he can fly away, and he will return to the origin point only a finite number of times.

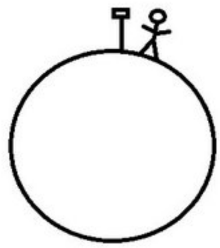


Figure 3: An example of the drunkard's walk

Now we consider fractal dimension. If $d \geq 2$, the fractal dimension of the set of points visited by a random walk is 2. The number of points of the lattice within distance r of the origin visited grows like r^2. Actually,

$$\#[points\ within\ distance\ r] \approx r^{fractal\ dimension}$$

The dimension of the path of the random walk does not depend on the dimension of the grid, assuming the grid has dimension at least two.

4 Self-avoiding Walks

We are going to change the game by adding some restrictions. One case is the *self-avoiding walk* (denoted as SAW). And several questions naturally come up.

- What happens if the walker does not return to any point she has already been to?
- How does this change the "dimension" of the set of points visited?

It actually depends a lot on how one specifies the problem. Certainly, it won't be interesting in one dimension, since if you cannot go backwards, it is pretty dull. So we usually talk about it in dimension 2 or higher. Let us consider all paths of length n that do not visit any point more than once, and make all such paths equally likely. This model was firstly proposed by Flory to model polymer chains.

4.1 How long is a typical self-avoiding walk?

The length of SAW is an example of an easily stated mathematical problem that is extremely difficult to analyze rigorously.

How much of an effect is there in requiring the past and future of a walk to avoid each other? Random walks have "dimension 2". In 3 dimension, two 2-dimensional

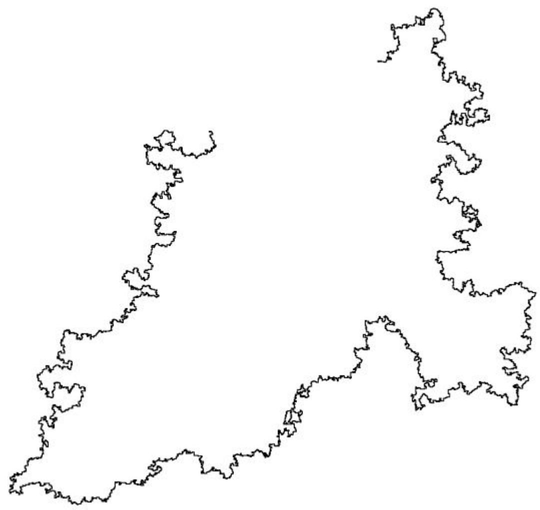


Figure 4: An example of SAW in 2 dimensions by V. Beffara

sets may hit each other. However, they do not intersect if they live in more than four dimensions. Thus, the self-avoidance constraint is not so important if $d \geq 5$. In this case, even if given self-avoidance, the distance is still of order $\sqrt{n}$, the same as the unrestricted random walk. This theorem is easy to get, but difficult to prove, and it eventually was proved by Slade and Hara.

In dimensions 2, 3, 4, the average length is not known rigorously. In fact, there are no nontrivial upper or lower bounds! However, there are conjectures.

Conjecture 1 (heuristic argument)**.** *(Flory) The average length of an n-step SAW is*

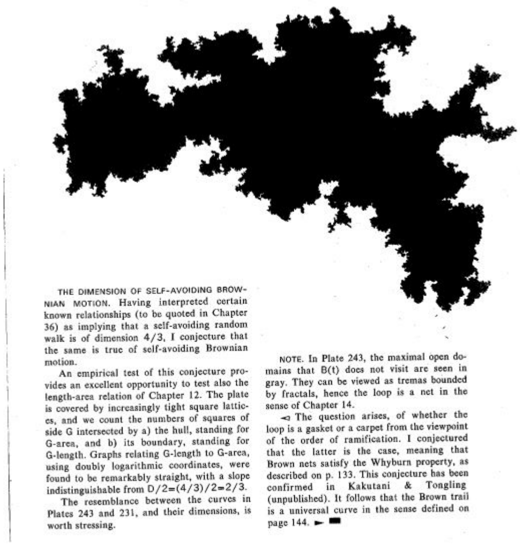
$$\begin{cases} n^{\frac{3}{d+2}}, & \text{if } d \leq 4 \\ n^{\frac{1}{2}}, & \text{if } d \geq 4. \end{cases}$$

This conjecture is correct for $d > 4$ and $d = 1$, which is particularly easy. For $d = 1$, the self-avoiding walk becomes a straight line, so the average length of a SAW is exactly n. For $d = 4$, it is expected to be right with "logarithmic corrections". Actually, this conjecture can be interpreted as *the fractal dimension of a SAW is* $\frac{d+2}{3}$.

If $d = 2$, the conjecture is expected to be true. In fact, there is a lot of good evidence for it:

- Very accurate numerical simulations;
- Much better heuristic arguments from theoretical physics using renormalization group and conformal field theory;
- Under certain assumptions of the limit, there is only one possible limit (a particular case of the Schramm-Loewner evolution). This limit object as fractal dimension $\frac{4}{3}$.

Despite this, there is no rigorous proof that the average length grows faster than $n^{\frac{1}{2}+\varepsilon}$ (it has to grow at least like $n^{\frac{1}{2}}$ since the walk has n distinct points) or that the average



THE DIMENSION OF SELF-AVOIDING BROWNIAN MOTION. Having interpreted certain known relationships (to be quoted in Chapter 36) as implying that a self-avoiding random walk is of dimension 4/3, I conjecture that the same is true of self-avoiding Brownian motion.

An empirical test of this conjecture provides an excellent opportunity to test also the length-area relation of Chapter 12. The plate is covered by increasingly tight square lattices, and we count the numbers of squares of side G intersected by a) the hull, standing for G-area, and b) its boundary, standing for G-length. Graphs relating G-length to G-area, using doubly logarithmic coordinates, were found to be remarkably straight, with a slope indistinguishable from D/2=(4/3)/2=2/3.

The resemblance between the curves in Plates 243 and 231, and their dimensions, is worth stressing.

NOTE. In Plate 243, the maximal open domains that B(t) does not visit are seen in gray. They can be viewed as tremas bounded by fractals, hence the loop is a net in the sense of Chapter 14.

◅ The question arises, of whether the loop is a gasket or a carpet from the viewpoint of the order of ramification. I conjectured that the latter is the case, meaning that Brown nets satisfy the Whyburn property, as described on p. 133. This conjecture has been confirmed in Kakutani & Tongling (unpublished). It follows that the Brown trail is a universal curve in the sense defined on page 144. ► ■

Figure 5: From geometry of Nature by Benoit Mandelbrot

length grows slower than $n^{1-\varepsilon}$ (the length is obviously no more than n.) If $d = 3$ (which is most important to polymer chemistry), the Flory prediction is $n^{\frac{3}{5}}$. While this is a nice number (and can be found in some chemistry text books), numerical simulations suggest that it is wrong. Actually, it should be $n^{.588\cdots}$.

Hence, there is no explicit conjecture for the number, maybe there never will be. However, the exponent is expected to be universal. SAWs on a different lattice should have the same exponent.

4.2 How many self-avoiding walks are there?

Let C_n be the number of SAWs of length n starting at the origin. There is a sub-multiplicative inequality which holds for C_n:

$$C_{n+m} \leq C_n C_m$$

There exists β called the connective constant such that $\lim_{n\to\infty} \frac{logC_n}{n} = \beta$. Roughly speaking, $C_n \approx \beta^n$. The value of β is not known and perhaps never will be. For now we only know that if $d = 2, \beta \approx 2.64$. The value of β in dimension higher than 2 is still an open question. In addition, the connective constant for the honeycomb lattice is known and it is $\sqrt{2+\sqrt{2}}$. This was found nonrigorously by Nienhuis and much later proved by Duminil-Copin and Smirnov.

There is an interesting thing mentioned in this colloquium which is the correction $\phi(n)$ to the leading behavior. Since C_n is approximately β^n, but we can write

$$C_n = \phi(n)\beta^n, \text{ and}$$
$$C_{2n} = C_n \cdot C_n \cdot p(n), \text{ where } p(n) = \frac{\phi(2n)}{\phi(n)^2}.$$

We view $p(n)$ as the probability that two SAWs starting at the same point have no other intersection. The behavior is conjectured to be universal (depends on the dimension but not the particular lattice).

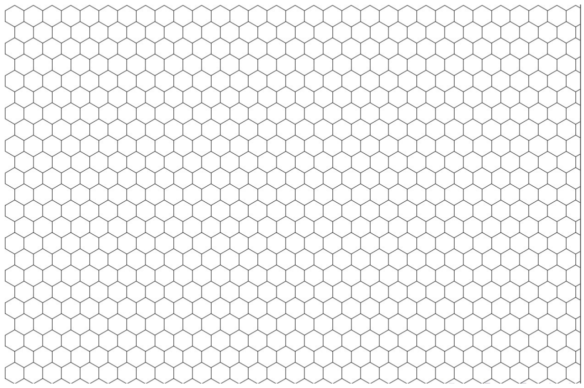


Figure 6: The honeycomb lattice

Then let us think about a question. Two simple random walkers starts at neighboring points. The first walker paints all sites she visits red. The second walker paints all he visits blue. After n steps, what is the chance that no site is painted both red and blue? The "fractal dimension" of the set of red points is 2, similarly for the blue points. Two 2-dimensional sets tend to intersect if the underlying space has fewer than four dimensions but do not tend to intersect there are more than four dimensions.

We summarize the above conclusion as follows:

- If $d \geq 5$, there is a positive probability that the red sites and the blue sites are completely separate.
- If $d \leq 4$, the set of points painted both red and blue is infinite. The "dimension" of the set of points painted both colors is $4 - d$ (the dimension of the intersection of two 2-dimensional sets in d dimensions).

4.3 A analogous problem about the points visited by two random walkers

Let $q(n)$ be the probability that no site has been painted more than one color after each walker has gone n steps. Then

$$q(n) \sim c \begin{cases} n^{-\zeta_d}, & d = 1, 2, 3 \\ (\log n)^{-\frac{1}{2}}, & d = 4 \end{cases}$$

We are going to analyze the formula in different dimension d. The key point is to find the parameter ζ_d.

- $\zeta_1 = 1$. In one dimension ζ_d is not difficult to find.
- $\zeta_2 = \frac{5}{8}$. In 2-dimension, ζ_2 has already been found, predicated by Duplantier and Kwon and proved by Lawler, Schramm, and Werner. The proof uses conformal invariance.
- The value of ζ_3 is unknown – in fact, it may well not be a rational number or any other "nice" number. Numerical simulation gives an estimate of $\zeta_3 \approx .29$.

In general, critical exponents for models in statistical physics are unknown above two dimensions and below the critical dimension (which is four in our case).

4.4 The same problem for two self-avoiding walkers

Let $p(n)$ be the probability that no site has been painted more than one color after each walker has gone n steps. Similarly, the formula for $p(n)$ is

$$p(n) \sim \begin{cases} n^{-\zeta_d}, & d = 1, 2, 3 \\ (\log n)^{-\frac{1}{4}}, & d = 4 \end{cases}$$

The parameter ζ_d in different dimensions are as follows:

- Nienhuis predicted $\zeta_2 = \frac{11}{32}$. This has been confirmed by Lawler, Schramm, and Werner for a related model, but is still open for SAW.
- The value of ζ_3 is unknown – in fact, it may well not be a rational number or any other "nice" number. Numerical simulation gives and estimate of $\zeta_3 \approx .16$.
- The $d = 4$ result is not proved, but there are related recent results of Brydges and Slade.

4.5 Loop-Erased Random Walk

Here we consider a new pattern of random walk, the loop-erased random walk. Get a path on self-avoiding walks by starting with a simple random walk and erasing the loops.

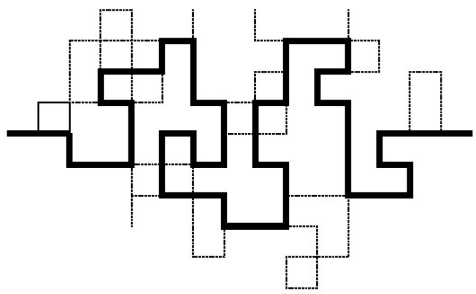


Figure 7: Loop-Erase Random Walk

In addition, we need to adjust the definition slightly in two dimensions(since random walk keeps returning).

Some questions about this new pattern of random walk come up.

- Does the path look like a self avoiding walk?
- What is the dimension of the path? For SAW it would be $\frac{4}{3}$ (if $d = 2$).

The loop-erased walk does not have the same fractal dimension as the self-avoiding walk. Consider the length of a walk of n steps. Then it has

$$\begin{gathered} n^{\frac{4}{5}},\ d = 2 \text{ (Kenyon)} \\ n^{.61\cdots},\ d = 3 \\ n^{\frac{1}{2}}(\log n)^{\frac{1}{6}},\ d = 4 \\ n^{\frac{1}{2}},\ d > 4 \end{gathered}$$

If $d = 2$, the fractal dimension is $\frac{5}{4}$.

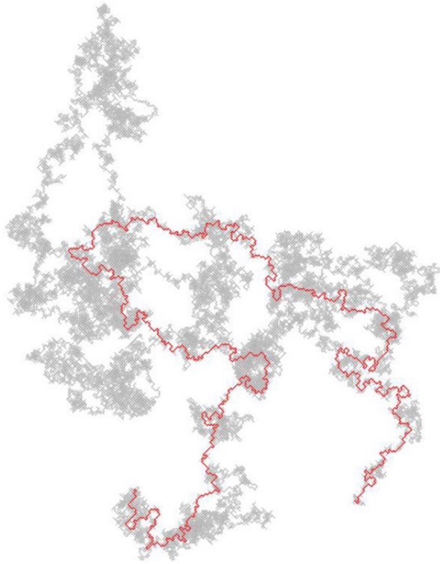


Figure 8: LERW by F. Viklund

5 Summary

There are many interesting random fractals that arise as limits of random walks. We have named only a few but there are many others coming from statistical physics. The behavior depends on the dimension. Below the "critical dimension" we get nontrivial "critical exponents". For $d = 2$, these exponents are often nice rational numbers. Analysis of this uses the conformal invariance properties of the limit. (Complex analysis). For $d = 3$, the exponents are universal in that they do not depend on the lattice (square vs. honeycomb,e.g.), but the values may never be known exactly. Many of the problems in the field are still open mathematical problems!