Self-avoiding walks and polygons: counting, joining and closing

arXiv:1504.05286

Alan Hammond
U.C. Berkeley

9th May, 2016

Self-avoiding walks

Self-avoiding walk is a fundamental example of a discrete model in statistical mechanics. It was introduced by Flory and Orr in the 1940s as a model in chemistry of a long chain of molecules.

A self-avoiding walk in $\mathbb{Z}^d$ of length n is

- a map $\gamma : \{0, \ldots, n\} \to \mathbb{Z}^d$
- that makes nearest-neighbor steps
- and visits no vertex twice.

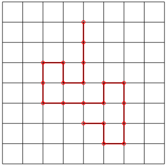


Figure: A planar self-avoiding walk of length twenty.

The uniform law on self-avoiding walks

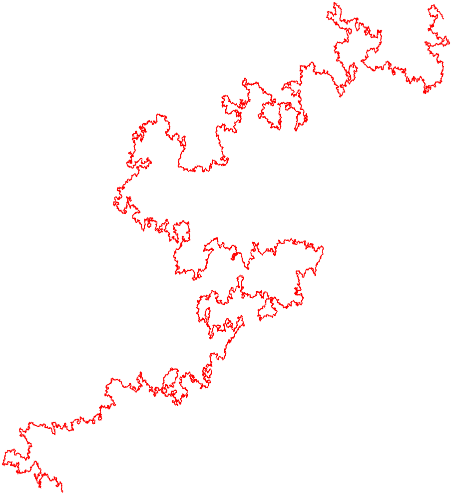
Let SAW_n denote the set of self-avoiding walks of length n that start at the origin.

Let W_n denote the uniform measure on SAW_n.

The length n walk under the law W_n will be denoted by Γ.

Simulation of planar SAW due to Tom Kennedy

SAW in plane - 1,000,000 steps



The endpoint of self-avoiding walk

Define the mean-squared displacement of the endpoint of the walk:

$$\langle ||\gamma_n||^2 \rangle = \frac{1}{|\mathrm{SAW}_n|} \sum_{\gamma \in \mathrm{SAW}_n} ||\gamma_n||^2 .$$

It is conjectured (and *rigorously known) that

$$\langle ||\gamma_n||^2 \rangle^{1/2} \; = \; n^{\nu + o(1)} \text{ where } \nu = \begin{cases} 1 & d = 1^* \\ 3/4 & d = 2 \quad \text{Nienhuis 1982} \\ \approx 0.59 & d = 3 \\ 1/2 & d = 4 \\ 1/2 & d \geq 5^* \text{ Hara, Slade 1992.} \end{cases}$$

Closing walks and self-avoiding polygons

A length n walk that ends at a location neighbouring the origin is said to *close*.

When such a walk γ closes, it is natural to add in the *missing edge* that connects γ_n and γ_0.

A self-avoiding polygon results.

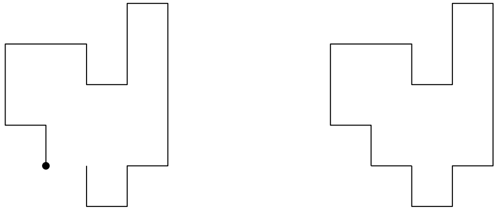


Figure: A closing walk and its polygon.

Counting walks and polygons

Let c_n denote the number of length n walks starting at the origin.

Let p_n denote the number of length n polygons *up to translation*.

Then the *closing probability* satisfies

$$W_n\big(\Gamma \text{ closes}\big) = \frac{2(n+1)p_{n+1}}{c_n}.$$

Walk subadditivity

A length $n + m$ walk γ can be severed at the vertex γ_n.

Two walks, of length n and m, result.

Thus, $c_{n+m} \leq c_n c_m$.

We may thus define the *connective* constant $\mu_W := \lim_{n \in \mathbb{N}} c_n^{1/n}$.

Polygon superadditivity

A polygon cannot be severed in two in this way.

However, a pair of polygons may be joined so that the new polygon's length equals the sum.

Thus, $p_{n+m} \geq \frac{1}{d-1} p_n p_m$.

We may thus define $\mu_P = \lim_{n \in 2\mathbb{N}} p_n^{1/n}$.

Polygon and walk deviation exponents

In fact, the two connective constants, μ_W and μ_P, are equal. This is because of a classic unfolding argument of Hammersley and Welsh.

Set μ to be the common value.

We have that $p_n \leq \mu^n \leq c_n$.

Let's set

$$p_n = n^{-\theta_n} \mu^n \quad \text{and} \quad c_n = n^{\xi_n} \mu^n \, .$$

Thus, θ_n and ξ_n are non-negative real numbers.

Hyperscaling relation

It is natural to define the polygon deviation exponent

$$\theta := \lim_{n \in 2\mathbb{N}} \theta_n$$

(though it may be very hard to prove that θ exists!)

A well known *hyperscaling* relation is believed to relate θ and ν:

$$\theta = d\nu + 1 \, .$$

We now present a heuristic derivation of the lower bound

$$\theta \geq 2\nu + 1$$

in two dimensions.

Hyperscaling relation lower bound

We will argue this in three steps:

- Step one: $\theta \geq \nu$;
- Step two: $\theta \geq \nu + 1$;
- Step three: $\theta \geq 2\nu + 1$.

Step one

This is Madras's 1995 polygon joining argument.

Take two polygons of length n.

There are order n^{ν} places where the second may be joined on the right to the first.

Thus,

$$p_{2n} \geq n^{\nu} p_n^2 \,,$$

and

$$p_n \leq n^{-\nu} \mu^n \,,$$

implying that $\theta \geq \nu$.

Step two

How can we progress from here?

We can join polygons in length pairs $(n+j, n-j)$, not just for $j = 0$ as before, but for all $|j| \leq n/2$.

We would seem to achieve

$$p_{2n} \geq n^{\nu} \sum_{j=-n/2}^{n/2} p_{n+j} p_{n-j} \, ,$$

and thus $\theta \geq \nu + 1$.

However, the joined polygons must have few macroscopic join points to reach this bound.

But it is plausible that they do typically.

Step three

We aim to move from $\theta \geq \nu + 1$ to $\theta \geq 2\nu + 1$.

All of the polygons we've been manufacturing are *double bubbles*.

We now argue that the fraction of length $2n$ polygons that are double bubbles is at most $Cn^{-\nu}$.

This provides the extra ν term that we seek.

Step three: escape from double bubble

Consider a typical length n polygon.

It crosses the strip $\big[-n^{\nu}, n^{\nu} \big] \times \mathbb{R}$ at least twice.

So there is a highest and a lowest crossing.

Now resample the uniform length n polygon by first sampling this law, and then forgetting about everything except:

- the highest crossing;
- and the lowest crossing, *up to vertical translation*.

Step three: escape from double bubble

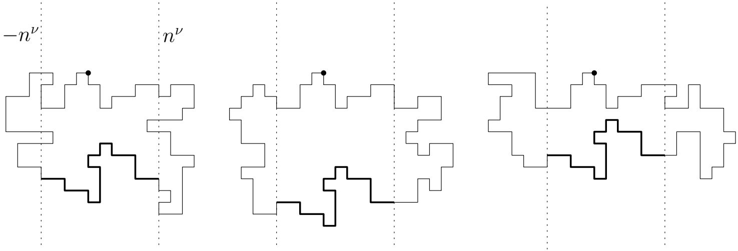


Figure: A uniform length n polygon on the left. Then two resamplings.

There's order n^ν vertical shifts that the lowermost crossing may undergo.

Only one of the them – the highest – leads to a double bubble.

So the chance of double bubble is at most $Cn^{-\nu}$.

Hyperscaling relation lower bound

So that's a non-rigorous argument for $\theta \geq 2\nu + 1$.

The derivation provides a useful framework for discussing rigorous proofs that use polygon joining.

Suppose a rigorous argument follows this three-step approach. Call it an (a, b, c)-argument, where these entries are the respective gains in θ made at each step.

So for example Madras' polygon joining argument is a $(1/2, 0, 0)$-argument.

The main results

The first result is a new lower bound on θ_n.

Recall that Madras' polygon joining shows that $\theta_n \geq 1/2$.

Theorem (1: Polygon Joining)

Let dimension $d = 2$. For a positive density subsequence, $\theta_n \geq 1$.

The main results

The next result concerns the closing probability $\mathsf{W}_n\big(\Gamma \,\mathrm{closes}\big)$.

It's not so obvious even that this quantity tends to zero in high n.

With Duminil-Copin, Glazman and Manolescu, we showed that

$$\mathsf{W}_n\big(\Gamma \,\mathrm{closes}\big) \leq n^{-1/4+o(1)} \,.$$

Theorem (2: Snake Method via Gaussian Pattern Fluctuation)

Consider any dimension d at least two. Then

$$\mathsf{W}_n\big(\Gamma \,\mathit{closes}\big) \leq n^{-1/2+o(1)} \,.$$

The main results

It's clear that this proof technique cannot do better than $n^{-1/2}$.

But we can push below $n^{-1/2}$ by mixing the two techniques – polygon joining and the snake method.

Theorem (3: Snake Method via Polygon Joining)

Let $d = 2$. Then, for a positive density subsequence of odd n,

$$\mathsf{W}_n\big(\Gamma \ closes\big) \leq n^{-6/11+o(1)} .$$

In fact, we may replace $6/11$ by $2/3$ conditionally *inter alia* on the existence of θ.

An overview of some aspects of the proofs

Theorem 1 – i.e., $\theta_n \geq 1$ on a subsequence – is derived by endeavouring to rework the three step derivation.

Madras already did step one rigorously – a $(1/2, 0, 0)$-argument.

To prove Theorem 1, we aim to implement step two – that is, to give a $(1/2, 1, 0)$-argument.

But we don't quite succeed, and wind up giving a $\big(1/2, 1 - 1/2, 0\big)$-argument.

An overview of the proof of Theorem 1

Remember that step two works out – and leads to a gain of one in the value of θ – if most double bubble polygons have few macroscopic join points.

In this rigorous version, we show only that there are typically at most $n^{1/2}$ such points.

An overview of the proof of Theorem 1

Why at most $n^{1/2}$ join points?

If there are more, then reflected walks may be modified to produce more than $e^{n^{1/2}}$ walks matched to each polygon.

But that contradicts the classical Hammersley-Welsh bound.

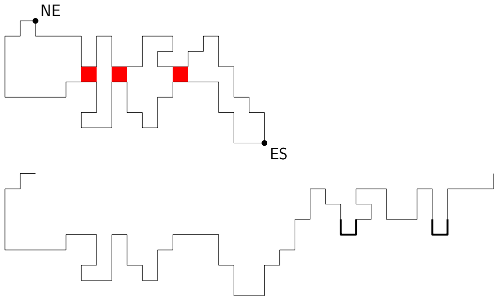


Figure: A polygon with its join points, then reflected and locally modified.

The snake method

To explain something of how Theorem 2 – closing probability is at most $n^{-1/2}$ – is obtained, we begin by discussing the snake method in a general guise.

It's a proof-by-contradiction technique for proving closing probability upper bounds.

It involves constructing sequences of laws of self-avoiding walks conditioned on increasingly severe avoidance constraints.

Explaining the snake method

First of all, a reflection argument shows that, for some $c > 0$,

$$\mathsf{W}_n\big(\Gamma \text{ closes}\big) \leq c\,.$$

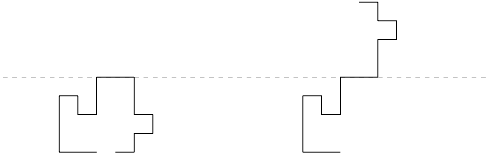


Figure: A closing walk may be reflected to form a non-closing alternative.

Explaining the snake method

How to improve this inference to show that

$$W_n\big(\Gamma \text{ closes}\big) \to 0\,?$$

Consider a typical *first part*.

Aim to argue that a walk in the half-space from the northeast corner typically meets the first part.

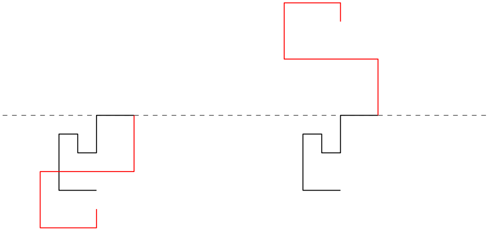


Figure: The reflection is viable even if the two parts meet on the left.

Explaining the snake method: polygonal invariance

To argue that the two half-space walks typically meet, *polygonal invariance* is an important tool.

Explaining the snake method: Kesten's pattern theorem

A pattern is any finite piece that may occur in the middle of a walk.

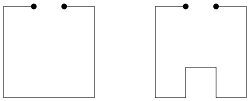


Figure: Type I and II patterns.

A classic result of Kesten asserts the ubiquity of any given pattern.

Theorem

For any pattern P, there exist $\delta \in (0,1)$ and $c > 0$ such that

$$\mathsf{W}_n\big(\, there\ are\ fewer\ than\ \delta n\ instances\ of\ P\ in\ \Gamma \big) \leq e^{-cn}\,.$$

Explaining the snake method: Kesten's pattern theorem

The snake method is a proof-by-contradiction technique.

Suppose that we're trying simply to show that

$$\mathsf{W}_n\big(\Gamma \text{ closes}\big) \to 0 \, .$$

Suppose instead that the closing probability is at least c.

Call a first part *charming* if the second part has positive probability to close the first.

Then for most $\ell \in [0, \ell]$, most length ℓ first parts are charming.

Explaining the snake method: general guise

But as the first part length ℓ rises, the second part length $n - \ell$ falls.

If we can show that the first part is often charming even if the second part length is not changing, then we have a powerful mechanism for manufacturing alternative walks by reflection.

Explaining the snake method: pattern fluctuation

Gaussian pattern fluctuation is a technique for showing that the second part length may remain fixed as the first part length ℓ varies.

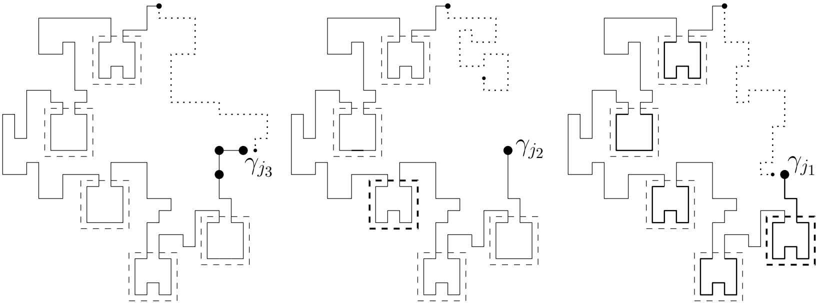


Figure: By switching a type *I* pattern in the first part to be of type *II*, two units of length accumulate in the first part.

Explaining the snake method: the $n^{-1/2}$ bound

Type *I* to type *II* pattern switching may be maintained for an order of $n^{1/2}$ steps without the law W_n noticing much.

This is in essence the same $n^{1/2}$ as in Theorem 2:

$$\mathsf{W}_n\big(\Gamma \text{ closes}\big) \leq n^{-1/2+o(1)} \, .$$

Explaining the snake method: beyond $n^{-1/2}$

How to push below the $n^{-1/2}$ barrier to reach Theorem 3:

$$\mathsf{W}_n\big(\Gamma \text{ closes}\big) \leq n^{-6/11+o(1)} \text{ subsequentially ?}$$

Use the snake method again. Not via pattern fluctuation but via *polygon joining*.

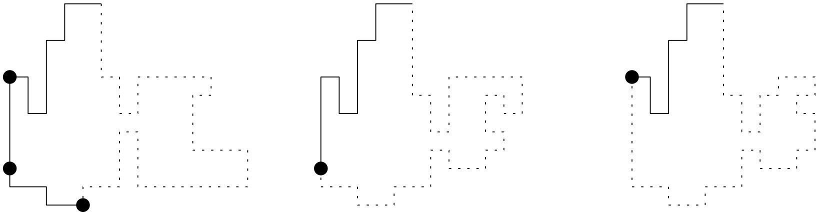


Figure: As the first part length falls, deflate the length of the right polygon in the join. The dotted second part remains of constant length.