A Geometric Perspective on Random Walks with Topological Constraints

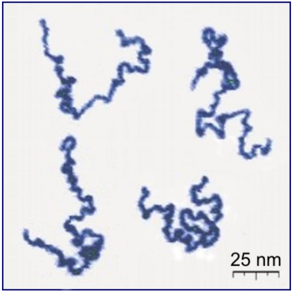
Clayton Shonkwiler

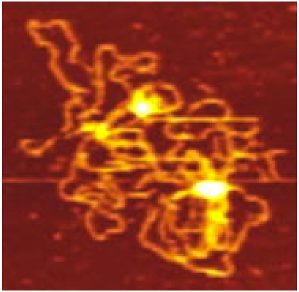
Colorado State University

LSU Graduate Student Colloquium
November 3, 2015

Statistical Physics Point of View

A polymer in solution takes on an ensemble of random shapes, with topology as the unique conserved quantity.



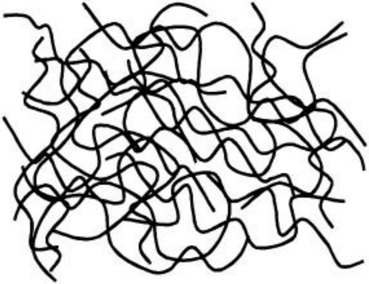


Protonated P2VP
Roiter/Minko
Clarkson University

Plasmid DNA
Alonso-Sarduy, Dietler Lab
EPF Lausanne

Statistical Physics Point of View

A polymer in solution takes on an ensemble of random shapes, with topology as the unique conserved quantity.



Schematic Image of Polymer Melt
Szamel Lab
CSU

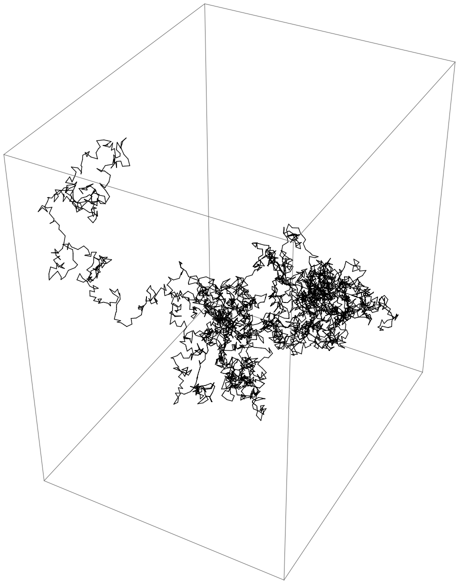
Statistical Physics Point of View

A polymer in solution takes on an ensemble of random shapes, with topology as the unique conserved quantity.

Physics Setup

> *Modern polymer physics is based on the analogy between a polymer chain and a random walk.*
> *—Alexander Grosberg, NYU.*

A Random Walk with 3,500 Steps



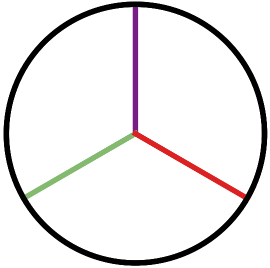
Ansatz

random walk $\iff$ *random point in some (nice!) moduli space*

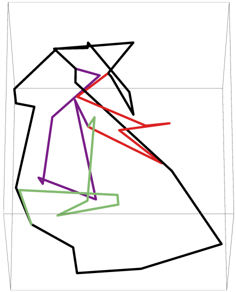
Scientific Idea

Use the (differential, symplectic, algebraic) geometry and topology of these moduli spaces to prove theorems and devise algorithms for studying random walks.

A **topologically constrained random walk** (TCRW) is a collection of random walks in $\mathbb{R}^3$ whose components are required to realize the edges of some fixed multigraph.

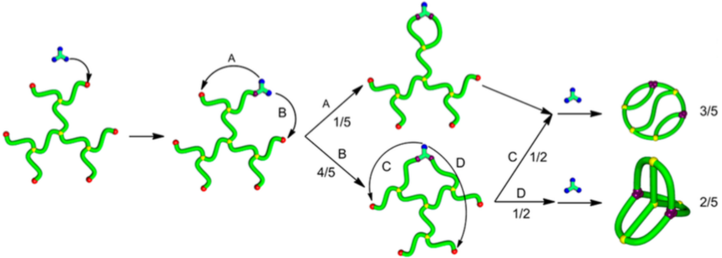


Abstract graph



TCRW

A **topologically constrained random walk** (TCRW) is a collection of random walks in $\mathbb{R}^3$ whose components are required to realize the edges of some fixed multigraph.



Tezuka Lab, Tokyo Institute of Technology

A synthetic $K_{3,3}$!

- What is the joint distribution of steps in a TCRW?
- What can we prove about TCRWs?
    - What is the joint distribution of vertex–vertex distances?
    - What is the expectation of radius of gyration?
    - Most common knot type among closed random walks?
- How do we sample TCRWs?

Closed Random Walks (a.k.a. Random Polygons)

The simplest multigraph with at least one edge is , which correponds to a *classical* random walk, modeling a *linear* polymer.

Closed Random Walks (a.k.a. Random Polygons)

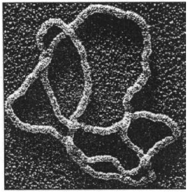
The simplest multigraph with at least one edge is , which correponds to a *classical* random walk, modeling a *linear* polymer.

The next simplest multigraph is , which yields a *closed* random walk (or *random polygon*), modeling a *ring* polymer.

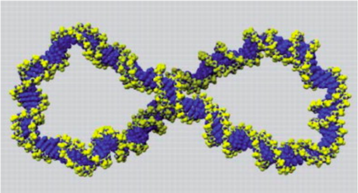
Closed Random Walks (a.k.a. Random Polygons)

The simplest multigraph with at least one edge is , which correponds to a *classical* random walk, modeling a *linear* polymer.

The next simplest multigraph is , which yields a *closed* random walk (or *random polygon*), modeling a *ring* polymer.



Knotted DNA
Wassermann et al.
Science **229**, 171–174

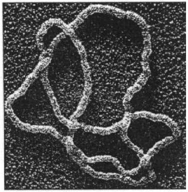


DNA Minicircle simulation
Harris Lab
University of Leeds, UK

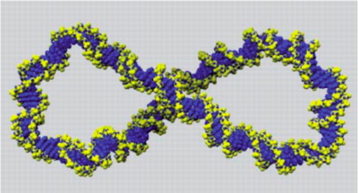
Closed Random Walks (a.k.a. Random Polygons)

The simplest multigraph with at least one edge is , which correponds to a *classical* random walk, modeling a *linear* polymer.

The next simplest multigraph is , which yields a *closed* random walk (or *random polygon*), modeling a *ring* polymer.



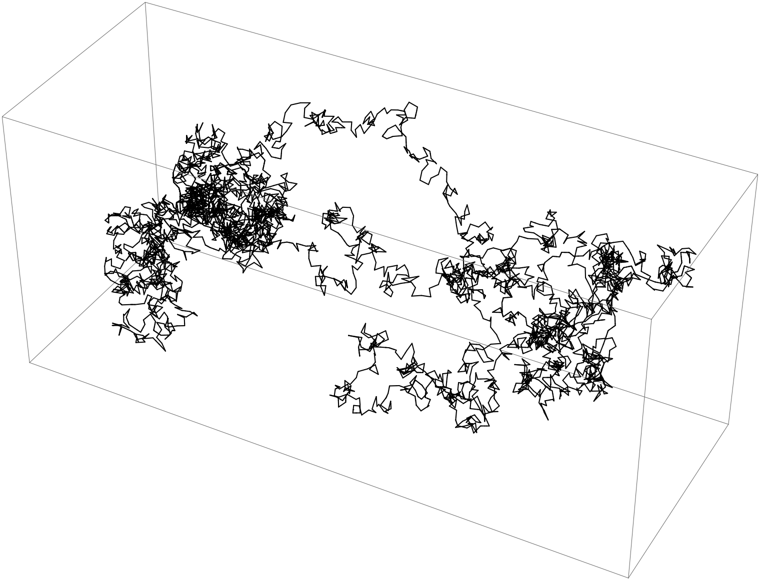
Knotted DNA
Wassermann et al.
Science **229**, 171–174



DNA Minicircle simulation
Harris Lab
University of Leeds, UK

We will focus on closed random walks in this talk.

A Closed Random Walk with 3,500 Steps



Definition

Plane polygonal arm P (up to translation) $\iff$ $\vec{w} \in \mathbb{C}^n$, where the $w_1, \ldots, w_i$ are the edge directions.

Lemma

If we write $w_i = z_i^2$, then P has length 1 $\iff$ $\vec{z} \in S^{2n-1} \subset \mathbb{C}^n$.

Proof.

$$\text{Length}(P) = \sum |w_i| = \sum |z_i|^2. \tag{1}$$

$\square$

Conclusion

(Open) planar n-gons of length 1 $\iff$ the complex sphere.

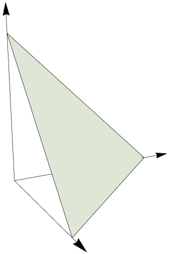
Question

Is this measure on polygonal arms a natural one?

Proposition (with Cantarella)

The sphere measure on open n-edge polygons of length one is equivalent to

- *choosing edge directions uniformly and independently*
- *choosing edge lengths uniformly on simplex*



$$= \{x_1, \ldots, x_n | x_i \geq 0, \sum x_i = 1\}$$

Second Construction: Plane Polygon Shapes

Lemma

Multiplying $\vec{z}$ by $e^{\mathbf{i}\theta}$ rotates P by 2θ.

Conclusion

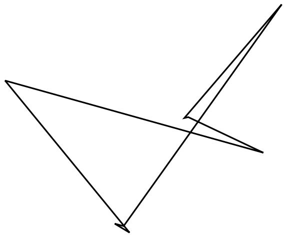
Planar n-gons (up to translations and rotation) $\iff$

$$\mathbb{C}P^n = S^{2n-1}/(\vec{z} \simeq e^{\mathbf{i}\theta}\vec{z})$$

This is already interesting, because it implies that a rotation and translation invariant distance between these shapes is given by measuring distance in $\mathbb{C}P^n$.

Construct real n-vectors $\vec{a} = \Re\vec{z}$, $\vec{b} = \Im\vec{z}$.

$$\begin{pmatrix} -0.18 & + & 0.05\mathbf{i} \\ -0.27 & + & -0.57\mathbf{i} \\ -0.73 & + & 0.1\mathbf{i} \\ -0.09 & + & -0.4\mathbf{i} \\ 0.01 & + & -0.09\mathbf{i} \\ -0.5 & + & -0.23\mathbf{i} \\ 0.34 & + & -0.66\mathbf{i} \end{pmatrix} \overset{\vec{z}}{\Longleftrightarrow}$$



Theorem (Hausmann/Knutson)

P closed and length 2 $\iff$ $\vec{a}, \vec{b}$ orthonormal.

Proof.

$z_i^2 = (a_i^2 - b_i^2) + (2a_i b_i)\mathbf{i}$, so

$$\sum z_i^2 = 0 \iff \sum a_i^2 = \sum b_i^2, \sum a_i b_i = 0.$$

Fourth Construction: Closed plane polygon shapes

Definition

The Stiefel manifold $V_k(\mathbb{R}^n) \iff$ orthonormal k-frames in $\mathbb{R}^n$.

Conclusion (Hausmann/Knutson)

Closed planar n-gons of length $2 \iff V_2(\mathbb{R}^n)$.

Lemma

Multiplying $\vec{z}$ *by* $e^{\mathbf{i}\theta}$ *rotates* $(\vec{a}, \vec{b})$ *in its own plane by* 2θ.

Definition

The Grassmann manifold $G_k(\mathbb{R}^n) \iff k$-planes in $\mathbb{R}^n$.

Conclusion (Hausmann/Knutson)

Closed planar n-gons of length 2 *(up to trans/rot)* $\iff G_2(\mathbb{R}^n)$.

Grassmannians and Stiefel manifolds are a sort of crossroads of mathematical fields. MathSciNet has $>$ 5,000 papers on the subject.

Proposition

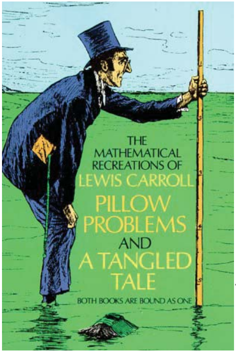
- $G_k(\mathbb{R}^n) = \frac{O(n)}{O(k)\times O(n-k)}$.
- $\dim G_k(\mathbb{R}^n) = k(n-k)$.
- $\mathrm{Vol}\, G_k(\mathbb{R}^n) = \frac{V(S^{n-1})\cdots V(S^{n-k})}{V(S^{k-1})\cdots V(S^1)}$
- *$G_k(\mathbb{R}^n)$ is homogeneous with transitive $O(n)$ action.*
- *There is a unique invariant (Haar) measure.*
- *Geometry and topology are very well known.*

Proposition

Dimension of closed, length 2, n-edge plane polygons $= 2(n-2) = 2n-4$.

Example: Triangle space

Lewis Carroll's Pillow Problem #58.



57. (25, 80)

In a given Triangle describe three Squares, whose bases shall lie along the sides of the Triangle, and whose upper edges shall form a Triangle;

(1) geometrically; (2) trigonometrically. [27/1/91

58. (25, 83)

Three Points are taken at random on an infinite Plane. Find the chance of their being the vertices of an obtuse-angled Triangle. [20/1/84

This problem proved difficult

> *The issue of choosing a "random triangle" is indeed problematic. I believe the difficulty is explained in large measure by the fact that there seems to be no natural group of transitive transformations acting on the set of triangles.*
>
> *–Stephen Portnoy, 1994*
> *(Editor, J. American Statistical Association)*

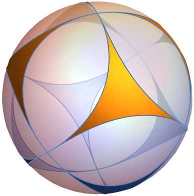
Observation

In the Grassmannian model, $O(3)$ is the natural group of geometric transformations on triangles $= G_2(\mathbb{R}^3)$.

A natural measure on triangles

Proposition (with Cantarella, Chapman, Needham)

If we lift $G_2(\mathbb{R}^3) = G_1(\mathbb{R}^3) = \mathbb{R}P^2$ to S^2, the measure is uniform:



(Gold region is acute triangles)
The fraction of obtuse triangles is

$$\frac{3}{2} - \frac{\log 8}{\pi} \simeq 83.8\%$$

Proposition (with Cantarella)

If we parametrize triangle space by edgelengths (assuming they sum to 2), the Grassmannian measure pushes forward to weighting each triangle by $1/\,\mathrm{Area}$.

Definition

The quaternions $\mathbb{H}$ are the skew-algebra over $\mathbb{R}$ defined by adding **i**, **j**, and **k** so that

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \quad \mathbf{ijk} = -1$$

Proposition

Unit quaternions (S^3) double-cover $\mathrm{SO}(3)$ *via the Hopf map.*

$$\mathrm{Hopf}(q) = (\bar{q}\mathbf{i}q, \bar{q}\mathbf{j}q, \bar{q}\mathbf{k}q),$$

where the entries turn out to be purely imaginary quaternions, and hence vectors in $\mathbb{R}^3$*.*

Fifth construction: (framed) space polygons

Definition

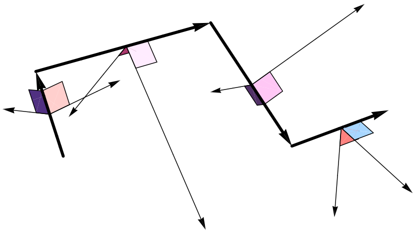
Framed n-gons in $\mathbb{R}^3 \iff$ vectors $\vec{w}$ in $C\,\mathrm{SO}(3)^n$.

Proposition

If we let $w_i = \mathrm{Hopf}(q_i)$*, then framed* n*-gons of total length 1* $\iff$ *unit sphere* $S^{4n-1} \subset \mathbb{H}^n$*.*

Proof.

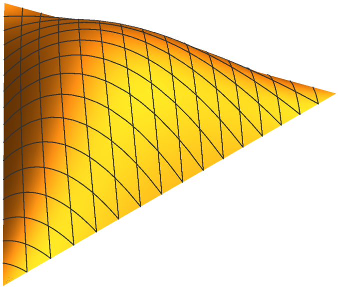
$|\bar{q}_i \mathbf{i} q_i| = |q_i|^2$, the edges of the polygon are $\bar{q}_i \mathbf{i} q_i$. $\square$



Proposition (with Cantarella)

The distribution of edges in the quaternionic model is:

- *directions are sampled independently, uniformly on* $(S^2)^n$.
- *lengths are sampled by the Dirichlet* $(2, \ldots, 2)$ *distribution on the simplex* $\{\vec{x} | x_i \geq 0, \sum x_i = 1\}$.



$\iff$ pdf is $\sim x_1 x_2 \cdots x_n$

Sixth construction: (framed) space polygon shapes

Proposition

Multiplying $\vec{q}$ by w rotates polygon by matrix $\operatorname{Hopf}(w) \in \mathrm{SO}(3)$.

Conclusion

Framed, length 1, space polygons (up to trans/rot) $\iff$

$$\mathbb{H}P^n = S^{4n-1}/(\vec{q} \simeq w\vec{q}, w \in \mathbb{H})$$

Again, this is already interesting, as the metric on $\mathbb{H}P^n$ then gives a translation and rotation invariant distance function for space polygons.

Seventh Construction: Closed framed space polygons

Every quaternion $q = a + b\mathbf{j}$, where $a, b \in \mathbb{C}$. This means that we can take *complex* vectors $(\vec{a}, \vec{b})$ corresponding to a quaternionic vector $\vec{q}$.

Proposition (Hausmann/Knutson)

P is closed, length 2 $\iff$ *the vectors* $(\vec{a}, \vec{b})$ *are Hermitian orthonormal.*

Proof.

$$\mathrm{Hopf}(a + b\mathbf{j}) = (\overline{a + b\mathbf{j}})\mathbf{i}(a + b\mathbf{j}) = \mathbf{i}(|a|^2 - |b|^2 + 2\bar{a}b\mathbf{j})$$

so we have

$$\sum \mathrm{Hopf}(a + b\mathbf{j}) = 0 \iff \sum |a|^2 = \sum |b^2|, \sum \bar{a}b = 0.$$

□

8th Cons.: Closed, rel. framed space poly shapes

Conclusion (Hausmann/Knutson)

Closed, framed space polygons $\iff$ $V_2(\mathbb{C}^n)$.

Proposition (Hausmann/Knutson)

The action of the matrix group $\mathrm{U}(2)$ *on* $V_2(C^n)$

- *rotates the polygon in space (*$\mathrm{SU}(2)$ *action)* ***and***
- *spins all vectors of the frame (*$\mathrm{U}(1)$ *action).*

Conclusion (Hausmann/Knutson)

Closed, rel. framed space polygons of length 2 $\iff$ $G_2(\mathbb{C}^n)$.

Again, there is a unique invariant (Haar) measure on $G_2(\mathbb{C}^n)$ which is a good candidate for the natural probability measure on closed (relatively framed) space polygons.

Idea

Translate closed random walk questions into questions about Haar measure on the complex Grassmannian of 2-planes, solve them there.

It's a natural principle that short arcs of long closed polygons should "look like" corresponding arcs of long open polygons.

Definition

Given two probability measures μ and ν on a measure space X, the *total variation* distance between μ and ν is

$$|\mu - \nu|_{\mathrm{TV}} = \max_{A \subset X} |\mu(A) - \nu(A)|$$

Theorem (Berglund)

The tv distance between k-edge arcs of open and closed n-edge framed space polygons is bounded by

$$2\left(\frac{4k+3}{4n-4k-3} + \frac{n^4}{(n-k-2)^4} - 1\right)$$

for large n, $< \frac{10k+17.5}{2n}$

Corollary (Berglund)

If f is a bounded function on k-edge arms,

$$\lim_{n\to\infty} \frac{|E(f, k\text{-edge arcs of } n\text{-edge closed polygons})|}{|E(f, k\text{-edge arcs of } n\text{-edge open polygons})|} \to 1$$

Proposition (with Cantarella, Grosberg, Kusner)

The expected value of total turning angle for an n-turn

- *open polygon is*
$$\frac{\pi}{2}n$$
- *closed polygon is*
$$\frac{\pi}{2}n + \frac{\pi}{4}\frac{2n}{2n-3}.$$

Corollary (with Cantarella, Grosberg, Kusner)

At least $1/3$ *of rel. framed hexagons and* $1/11$ *of rel. framed heptagons are unknots.*

Corollary (with Cantarella, Grosberg, Kusner)

At least $1/3$ *of rel. framed hexagons and* $1/11$ *of rel. framed heptagons are unknots.*

Proof.

Let x be the fraction of n-gons with total curvature greater than 4π (by the Fáry-Milnor theorem, these are the only polygons which may be knotted). The expected value of total curvature then satisfies

$$E(\kappa) > 4\pi x + 2\pi(1 - x).$$

Solving for x and using our total curvature expectation, we see that

$$x < \frac{(n-2)(n-3)}{2(2n-3)}.$$

$\square$

Conjecture (Frisch–Wassermann, Delbrück, 1960s)

As $n \to \infty$, the probability that an n-gon is unknotted is

$$P(unknot) < e^{-\alpha n}$$

for some $\alpha > 0$.

Conjecture (Frisch–Wassermann, Delbrück, 1960s)

As $n \to \infty$, the probability that an n-gon is unknotted is

$$P(unknot) < e^{-\alpha n}$$

for some $\alpha > 0$.

Proved in different random polygon models by Sumners–Whittington, Pippinger, and Diao in the 1980s-90s.

Conjecture (Frisch–Wassermann, Delbrück, 1960s)
As $n \to \infty$, the probability that an n-gon is unknotted is

$$P(unknot) < e^{-\alpha n}$$

for some $\alpha > 0$.

Proved in different random polygon models by Sumners–Whittington, Pippinger, and Diao in the 1980s-90s.

...but how big is α? Nobody knows.

Sampling random polygons in quaternionic model

Theorem (with Cantarella, Deguchi)

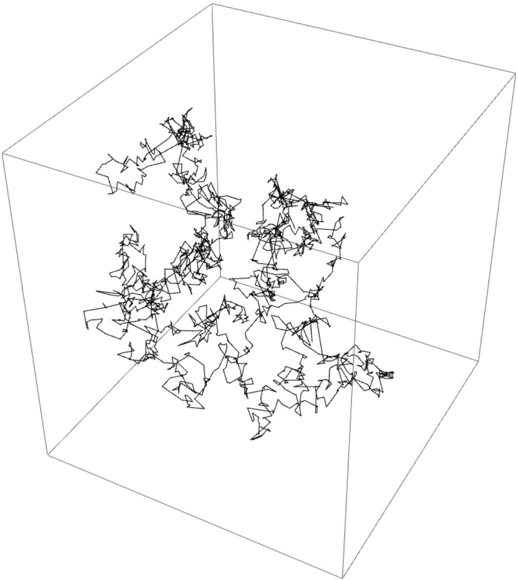
We can randomly generate framed n-gons uniformly with respect to the symmetric measure in $O(n)$ time.

In[9]:=
```
RandomComplexVector[n_] := Apply[Complex,
    Partition[#, 2] & /@ RandomVariate[NormalDistribution[], {1, 2 n}], {2}][[1]];

ComplexDot[A_, B_] := Dot[A, Conjugate[B]];
ComplexNormalize[A_] := (1 / Sqrt[Re[ComplexDot[A, A]]]) A;

RandomComplexFrame[n_] := Module[{a, b, A, B},
   {a, b} = {RandomComplexVector[n], RandomComplexVector[n]};
    A = ComplexNormalize[a];
    B = ComplexNormalize[b - Conjugate[ComplexDot[A, b]] A];
   {A, B}
  ];
```

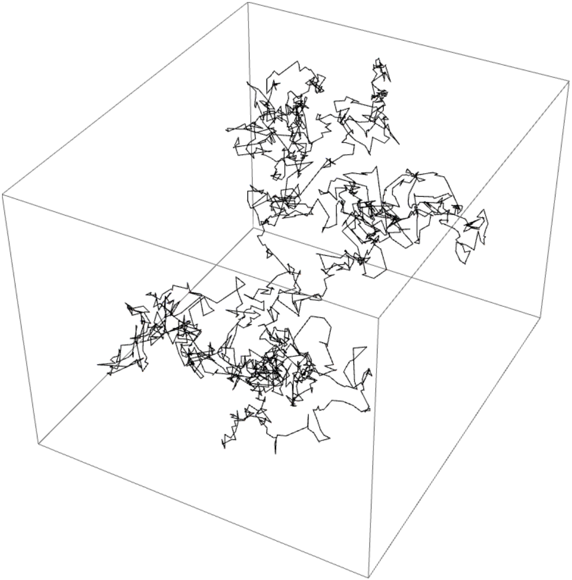
Now we need only apply the Hopf map to generate an edge set:

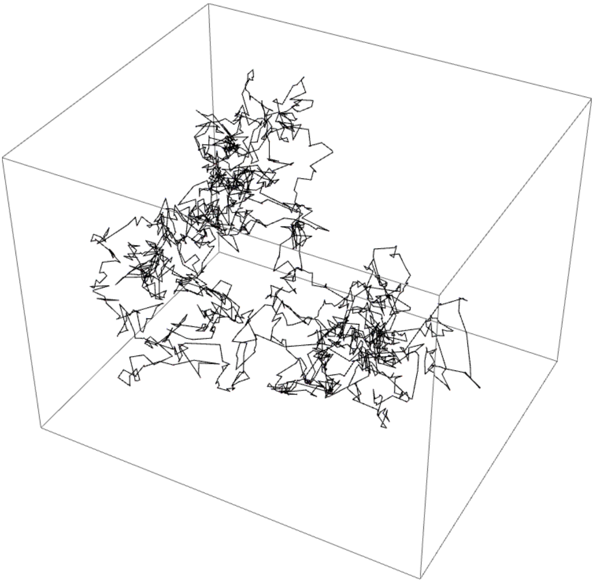
In[6]:=
```
ToEdges[{A_, B_}] := {#[[2]], #[[3]], #[[4]]} & /@ (HopfMap /@ Transpose[{A, B}]);
```

Random 2,000-gons

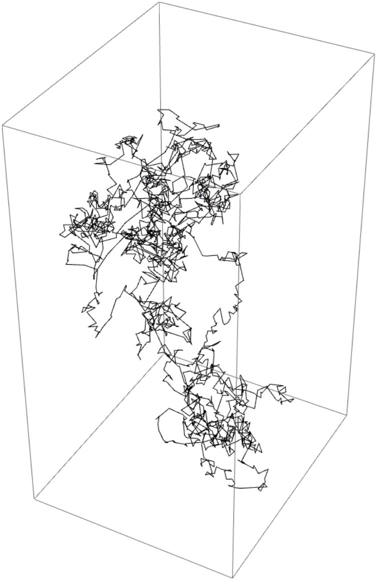


Random 2,000-gons

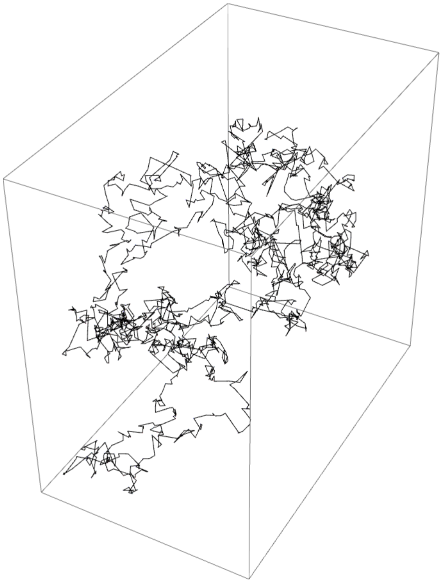




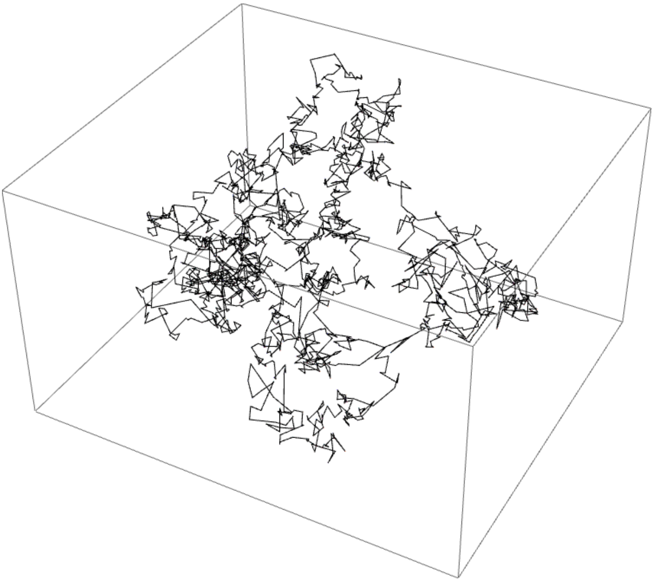
Random 2,000-gons



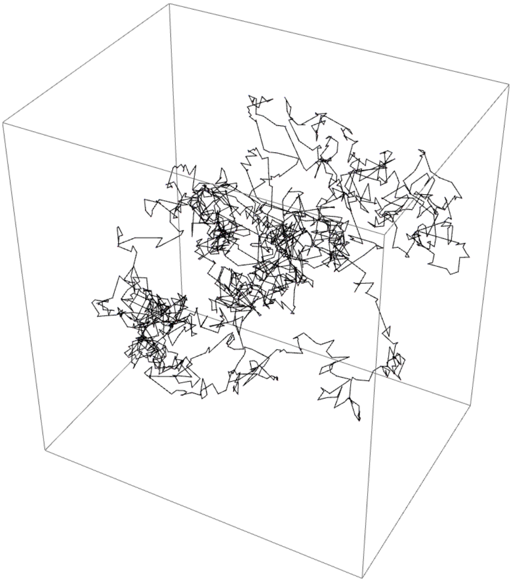
Random 2,000-gons



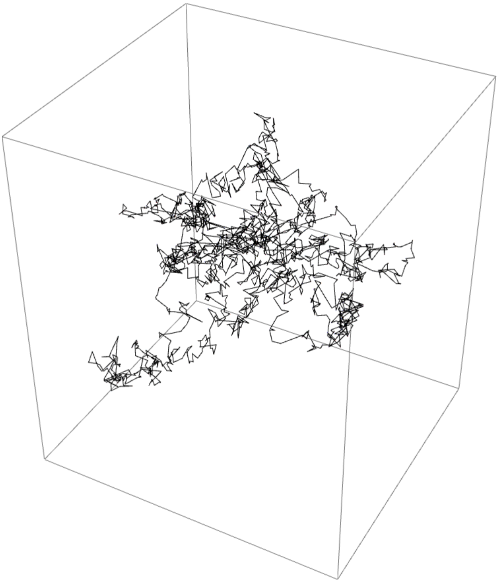
Random 2,000-gons



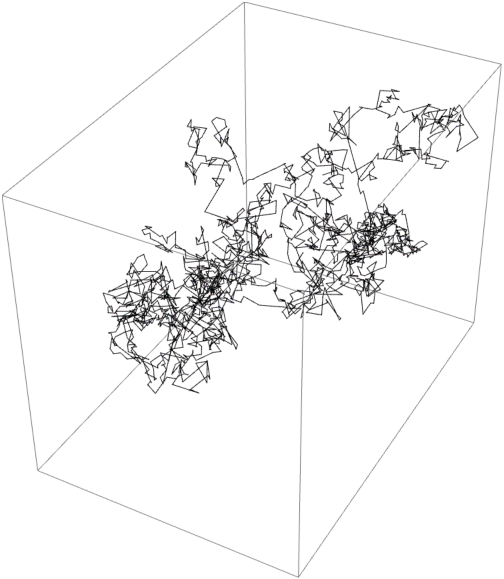
Random 2,000-gons



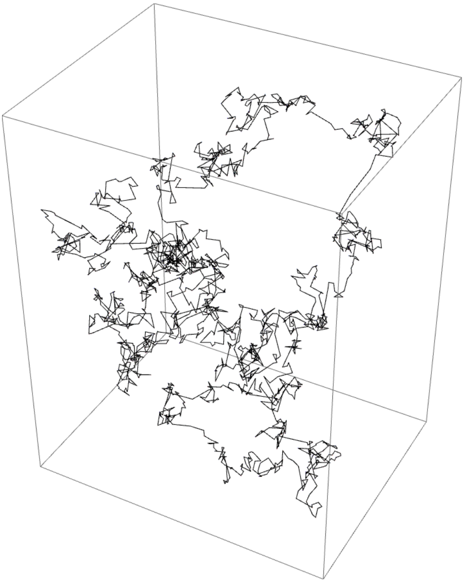
Random 2,000-gons



Random 2,000-gons



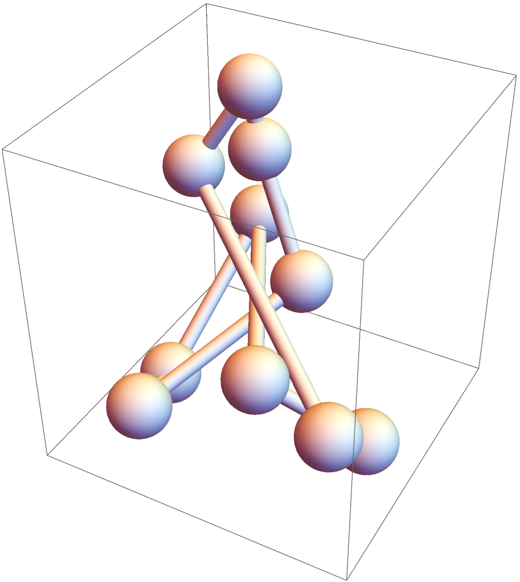
Random 2,000-gons



Open Questions

A few open questions to think about:

- TCRWs based on more complicated graphs.
- self-avoiding random walks
- a theoretical understanding of knotting in these models
- a general theory of random piecewise-linear submanifolds?



Observation

If vertices v_0, v_k of P collide, then the first k edges and the last $n - k$ edges of the polygon form smaller polygons. So the Grassmannian representation of P is contained in

$$G_2(\mathbb{C}^k) \times G_2(\mathbb{C}^{n-k}) \subset G_2(\mathbb{C}^n)$$

Conclusion

Polygons that avoid balls around vertices $\iff$

complement of union of neighborhoods of $G_2(\mathbb{C}^k) \times G_2(\mathbb{C}^{n-k})$

Question

What can we say about this complement?

Thank you for listening!

- *Probability Theory of Random Polygons from the Quaternionic Viewpoint*
  Jason Cantarella, Tetsuo Deguchi, and Clayton Shonkwiler
  Communications on Pure and Applied Mathematics **67** (2014), no. 10, 658–1699.

- *The Expected Total Curvature of Random Polygons*
  Jason Cantarella, Alexander Y Grosberg, Robert Kusner, and Clayton Shonkwiler
  American Journal of Mathematics **137** (2015), no. 2, 411–438

`http://arxiv.org/a/shonkwiler_c_1`