A Geometric Perspective on Random Walks with Topological Constraints

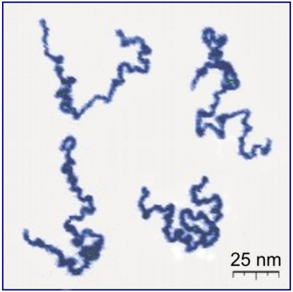
Clayton Shonkwiler

Colorado State University

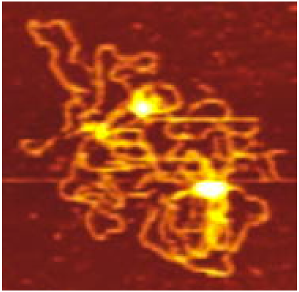
Wake Forest University
September 9, 2015

Statistical Physics Point of View

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



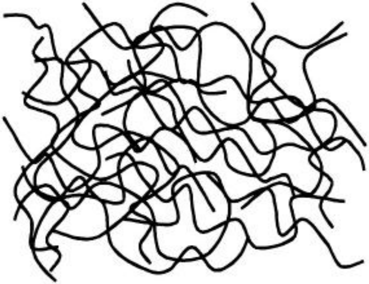
Protonated P2VP
Roiter/Minko
Clarkson University



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

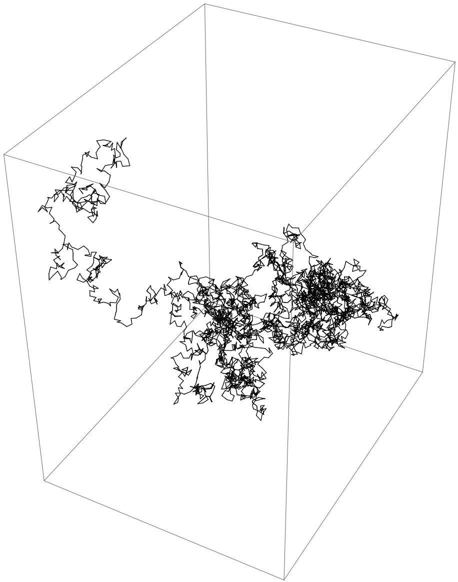
Statistical Physics Point of View

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

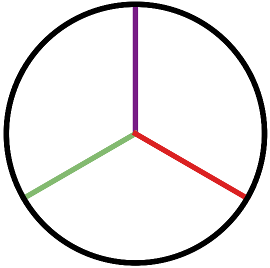
Physics Setup

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

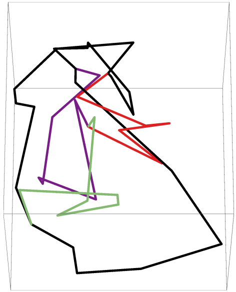
A Random Walk with 3,500 Steps



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.



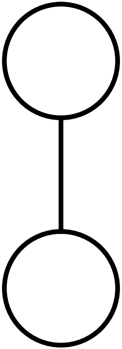
Abstract graph



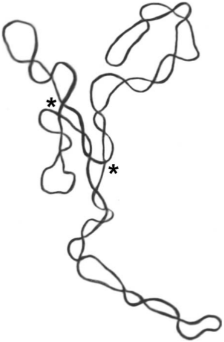
TCRW

Topologically Constrained 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**.**



Barbell graph

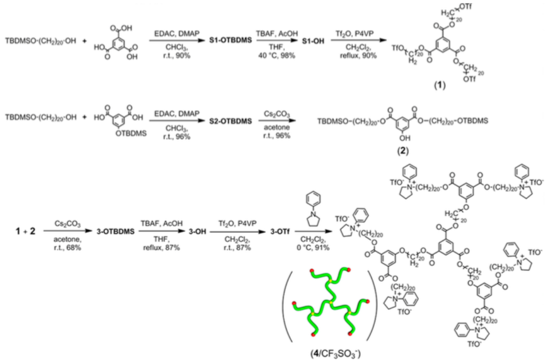


Branched DNA
Benham–Mielke *Ann. Rev. Biomed. Eng.* **7**, 21–53

Topologically Constrained Random Walks

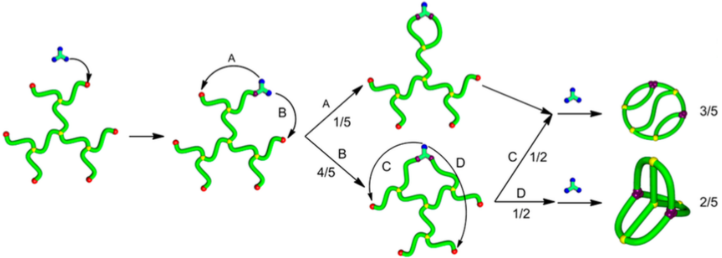
A **topologically constrained random walk** (TCRW) is a collection of random walks in $\mathbb{R}^3$ whose components are required to realize the edges of some fixed multigraph.

Scheme 1. Synthesis of a Hexafunctional Dendritic Precursor, $4/CF_3SO_3^-$



Tezuka Lab, Tokyo Institute of Technology

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 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?
  - What is the expectation of total turning?

- 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?
  - What is the expectation of total turning?
- How do we sample TCRWs?

- 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?
    - What is the expectation of total turning?
- How do we sample TCRWs?

Point of Talk

We can use geometric understanding of the moduli space of TCRWs based on a fixed graph to answer these questions.

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)

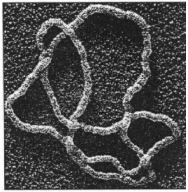
The simplest multigraph with at least one edge is , which correponds to a *classical* random walk, modeling a *linear* polymer.

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

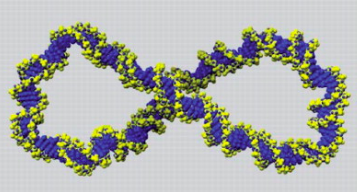
Closed Random Walks (a.k.a. Random Polygons)

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

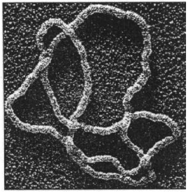


DNA Minicircle simulation
Harris Lab
University of Leeds, UK

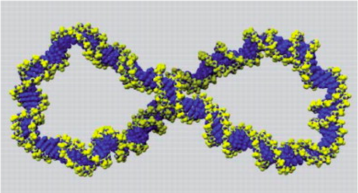
Closed Random Walks (a.k.a. Random Polygons)

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.



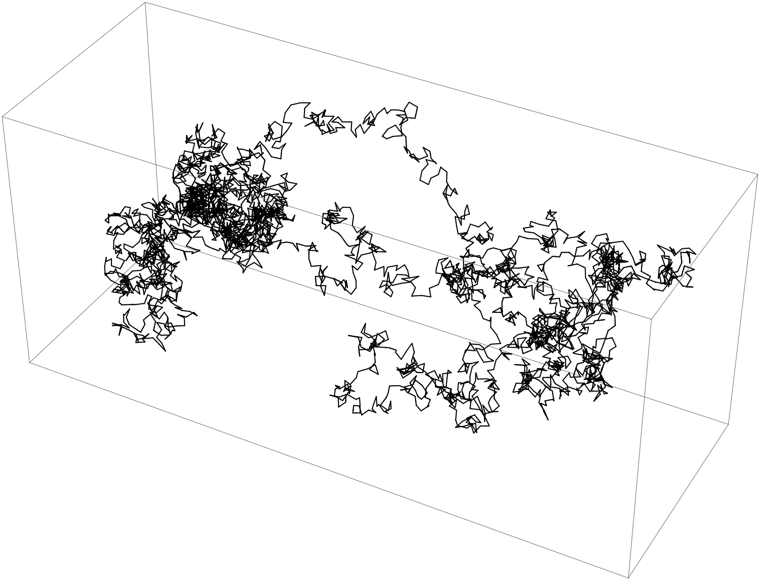
Knotted DNA
Wassermann et al.
Science **229**, 171–174



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



(Incomplete?) History of Sampling Algorithms

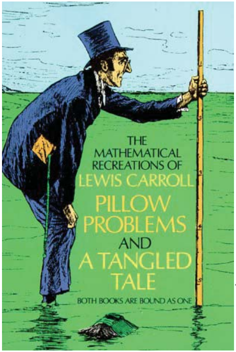
- Markov Chain Algorithms
  - crankshaft (Vologoskii 1979, Klenin 1988)
  - polygonal fold (Millett 1994)
- Direct Sampling Algorithms
  - triangle method (Moore 2004)
  - generalized hedgehog method (Varela 2009)
  - sinc integral method (Moore 2005, Diao 2011)

(Incomplete?) History of Sampling Algorithms

- Markov Chain Algorithms
  - crankshaft (Vologoskii et al. 1979, Klenin et al. 1988)
    - convergence to correct distribution unproved
  - polygonal fold (Millett 1994)
    - convergence to correct distribution unproved
- Direct Sampling Algorithms
  - triangle method (Moore et al. 2004)
    - samples a subset of closed polygons
  - generalized hedgehog method (Varela et al. 2009)
    - unproved whether this is correct distribution
  - sinc integral method (Moore et al. 2005, Diao et al. 2011)
    - requires sampling complicated 1-d polynomial densities

A pillow problem

A side benefit: if we can understand random polygons, we can answer Lewis Carroll's Pillow Problem #58.



14 *PILLOW-PROBLEMS.*

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

What does it mean to take a random triangle?

Question

What does it mean to choose a random triangle?

What does it mean to take a random triangle?

Question

What does it mean to choose a random triangle?

Statistician's Answer

> *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)*

What does it mean to take a random triangle?

Question

What does it mean to choose a random triangle?

What does it mean to take a random triangle?

Question

What does it mean to choose a random triangle?

Applied Mathematician's Answer

> *We will add a purely geometrical derivation of the picture of triangle space, delve into the linear algebra point of view, and connect triangles to random matrix theory.*
>
> *–Alan Edelman and Gil Strang, 2012*

What does it mean to take a random triangle?

Question

What does it mean to choose a random triangle?

What does it mean to take a random triangle?

Question

What does it mean to choose a random triangle?

Differential Geometer's Answer

> *Pick by the measure defined by the volume form of the natural Riemannian metric on the manifold of 3-gons, of course. That ought to be a special case of the manifold of n-gons.*
>
> *But what's the manifold of n-gons?*
>
> *And what makes one metric natural?*
>
> *Don't algebraic geometers understand this?*
>
> *–Jason Cantarella to me, a few years ago*

Lead-Up to the Algebraic Geometer's Answer

(Seemingly) Trivial Observation 1

Given $z_1 = a_1 + \mathbf{i}b_1, \ldots, z_n = a_n + \mathbf{i}b_n$

$$\begin{aligned}
\sum z_i^2 &= (a_1^2 - b_1^2) + \mathbf{i}(2a_1b_1) + \cdots + (a_n^2 - b_n^2) + \mathbf{i}(2a_nb_n) \\
&= (a_1^2 + \cdots + a_n^2) - (b_1^2 + \cdots + b_n^2) + \mathbf{i}\,2(a_1b_1 + \cdots + a_nb_n) \\
&= (|\vec{a}|^2 - |\vec{b}|^2) + \mathbf{i}\,2\left\langle \vec{a}, \vec{b} \right\rangle
\end{aligned}$$

(Seemingly) Trivial Observation 1

Given $z_1 = a_1 + \mathbf{i}b_1, \ldots, z_n = a_n + \mathbf{i}b_n$

$$\begin{aligned}
\sum z_i^2 &= (a_1^2 - b_1^2) + \mathbf{i}(2a_1 b_1) + \cdots + (a_n^2 - b_n^2) + \mathbf{i}(2a_n b_n) \\
&= (a_1^2 + \cdots + a_n^2) - (b_1^2 + \cdots + b_n^2) + \mathbf{i}\,2(a_1 b_1 + \cdots + a_n b_n) \\
&= (|\vec{a}|^2 - |\vec{b}|^2) + \mathbf{i}\,2\left\langle \vec{a}, \vec{b} \right\rangle
\end{aligned}$$

(Seemingly) Trivial Observation 2

Given $z_1 = a_1 + \mathbf{i}b_1, \ldots, z_n = a_n + \mathbf{i}b_n$

$$\begin{aligned}
|z_1^2| + \cdots + |z_n^2| &= |z_1|^2 + \cdots + |z_n|^2 \\
&= (a_1^2 + b_1^2) + \cdots + (a_n^2 + b_n^2) \\
&= (a_1^2 + \cdots + a_n^2) + (b_1^2 + \cdots + b_n^2) \\
&= |\vec{a}|^2 + |\vec{b}|^2
\end{aligned}$$

The Algebraic Geometer's Answer

Theorem (Hausmann and Knutson, 1997)

The space of closed planar n-gons with length 2, up to translation and rotation, is identified with the Grassmann manifold $G_2(\mathbb{R}^n)$ of 2-planes in $\mathbb{R}^n$.

The Algebraic Geometer's Answer

Theorem (Hausmann and Knutson, 1997)

The space of closed planar n-gons with length 2, up to translation and rotation, is identified with the Grassmann manifold $G_2(\mathbb{R}^n)$ of 2-planes in $\mathbb{R}^n$.

Proof.

Take an orthonormal frame $\vec{a}, \vec{b}$ for the plane, let

$$\vec{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} a_1 + \mathbf{i} b_1 \\ \vdots \\ a_n + \mathbf{i} b_n \end{pmatrix}, \quad v_\ell = \sum_{j=1}^{\ell} z_j^2$$

By the observations, $v_0 = v_n = 0$, $\sum |v_{\ell+1} - v_\ell| = 2$. Rotating the frame a, b in their plane rotates the polygon in the complex plane. $\square$

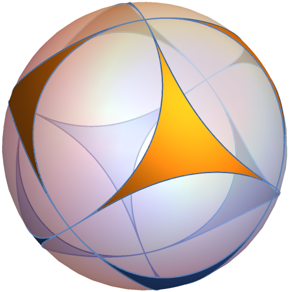
Theorem (with Cantarella and Deguchi)

The volume form of the standard Riemannian metric on $G_2(\mathbb{R}^n)$ – which is to say, Haar measure – defines the natural probability measure on closed, planar n-gons of length 2 up to translation and rotation. It has a (transitive) action by isometries given by the action of $SO(n)$ on $G_2(\mathbb{R}^n)$.

So *random triangles are points selected uniformly on $\mathbb{RP}^2$* since

$$\text{random triangle} \to \text{random point in } G_2(\mathbb{R}^3) \cong G_1(\mathbb{R}^3) = \mathbb{RP}^2.$$

Putting the pillow problem to bed



Acute triangles (gold) turn out to be defined by natural algebraic conditions on $\mathbb{RP}^2$ (or the sphere).

Proposition (with Cantarella, Chapman, and Needham, 2015)

The fraction of obtuse triangles is

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

Theorem (with Cantarella and Deguchi)

The volume form of the standard Riemannian metric on $G_2(\mathbb{C}^n)$ defines the natural probability measure on closed space n-gons of length 2 up to translation and rotation. It has a (transitive) action by isometries given by the action of $U(n)$ on $G_2(\mathbb{C}^n)$.

Proof.

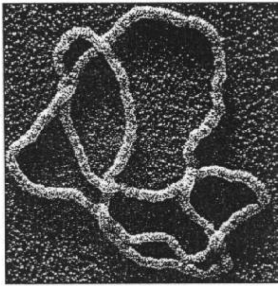
Again, we use an identification due to Hausmann and Knutson where

- instead of combining two real vectors to make one complex vector, we combine two complex vectors to get one quaternionic vector
- instead of squaring complex numbers, we apply the Hopf map to quaternions

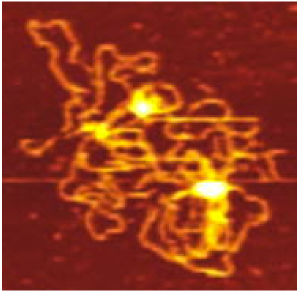
$\square$

Statistical Physics Point of View

A ring polymer in solution takes on an ensemble of random shapes, with topology (knot type!) as the unique conserved quantity.



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



Plasmid DNA
Alonso-Sarduy, Dietler Lab
EPF Lausanne

Definition

The total curvature of a space polygon is the sum of its turning angles.

Definition

The total curvature of a space polygon is the sum of its turning angles.

Theorem (with Cantarella, Grosberg, and Kusner)

The expected total curvature of a random n-gon of length 2 sampled according to the measure on $G_2(\mathbb{C}^n)$ is

$$\frac{\pi}{2}n + \frac{\pi}{4}\frac{2n}{2n-3}.$$

Definition

The total curvature of a space polygon is the sum of its turning angles.

Theorem (with Cantarella, Grosberg, and Kusner)

The expected total curvature of a random n-gon of length 2 sampled according to the measure on $G_2(\mathbb{C}^n)$ is

$$\frac{\pi}{2} n + \frac{\pi}{4} \frac{2n}{2n-3}.$$

Corollary (with Cantarella, Grosberg, and Kusner)

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

How can we sample and determine distributions of knot types?

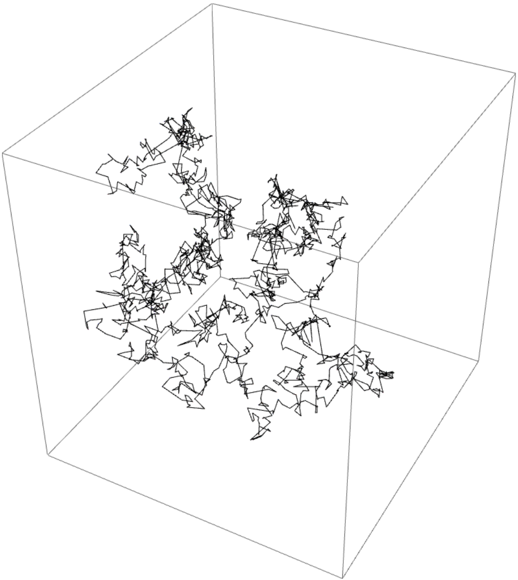
Proposition (classical?)

The natural measure on $G_2(\mathbb{C}^n)$ is obtained by generating two random complex n-vectors with independent Gaussian coordinates and their span.

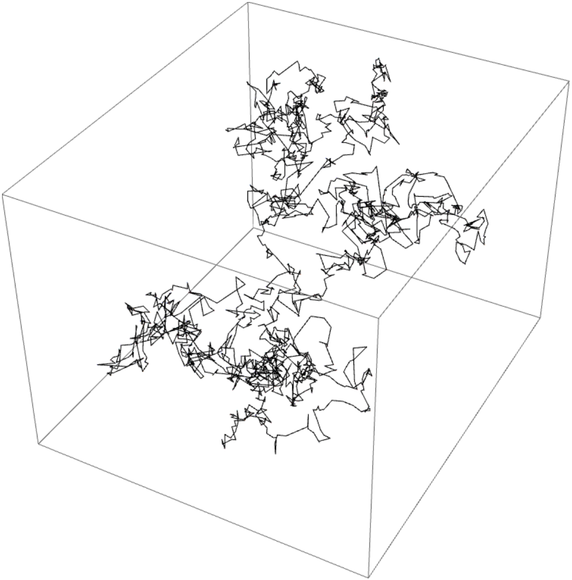
```
In[9]:= RandomComplexVector[n_] := Apply[Complex,
          Partition[#, 2] & /@ RandomVariate[NormalDistribution[], {1, 2 n}], {2}][[1]];

        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}
          ];
```

Using this, we can generate ensembles of random polygons . . .

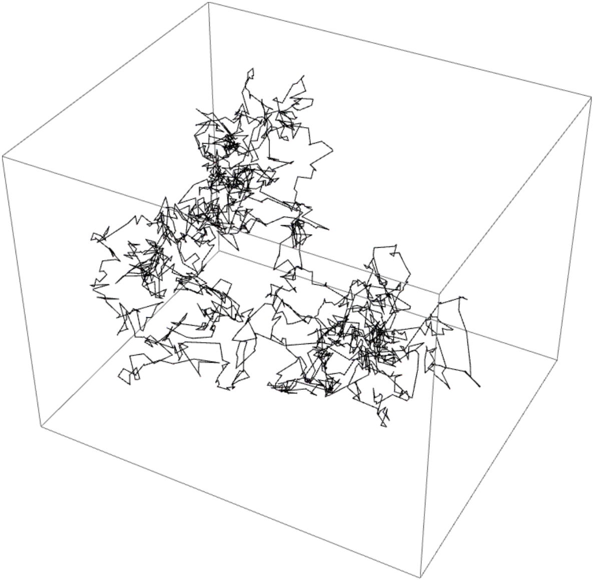
Random 2,000-gons



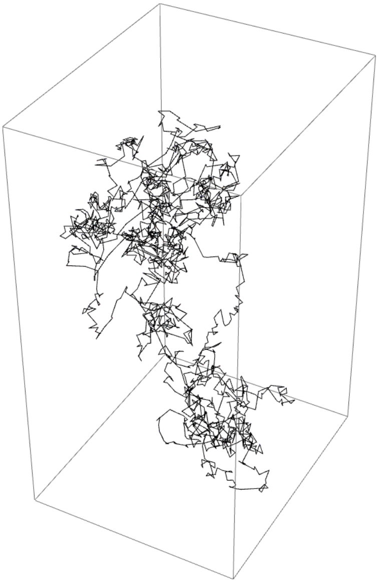
Random 2,000-gons



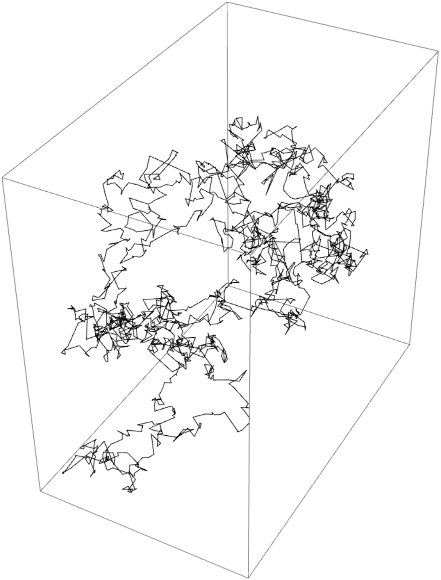
Random 2,000-gons



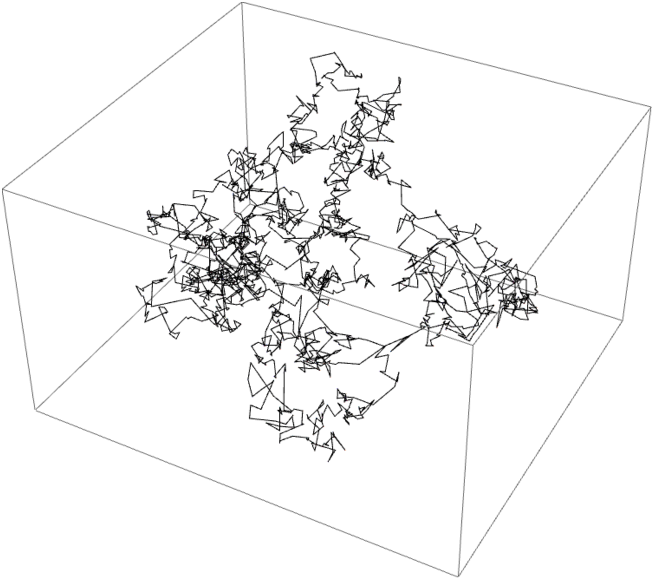
Random 2,000-gons



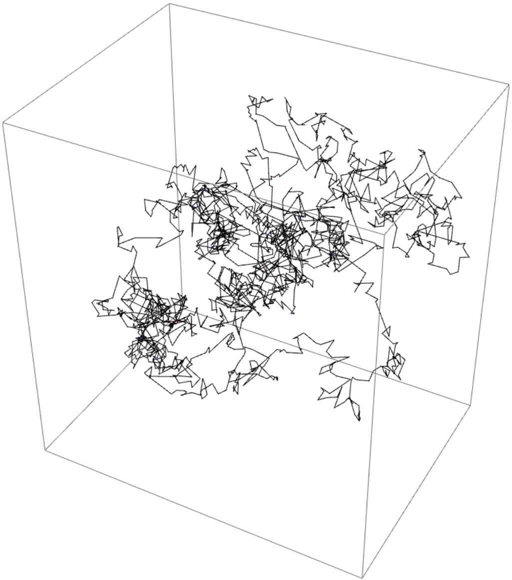
Random 2,000-gons



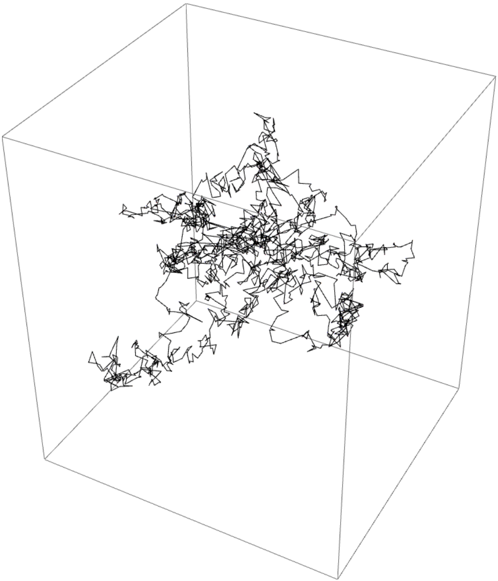
Random 2,000-gons



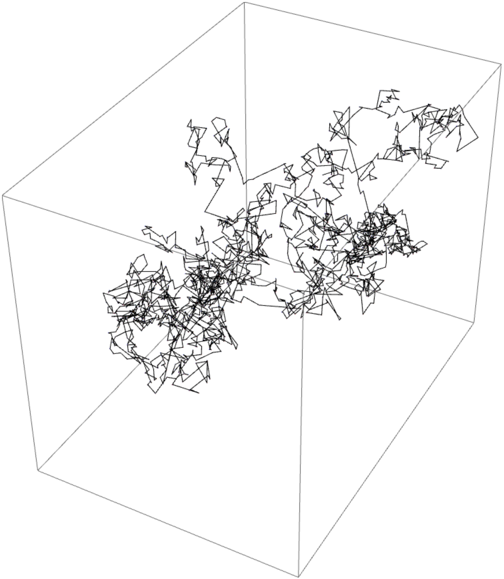
Random 2,000-gons



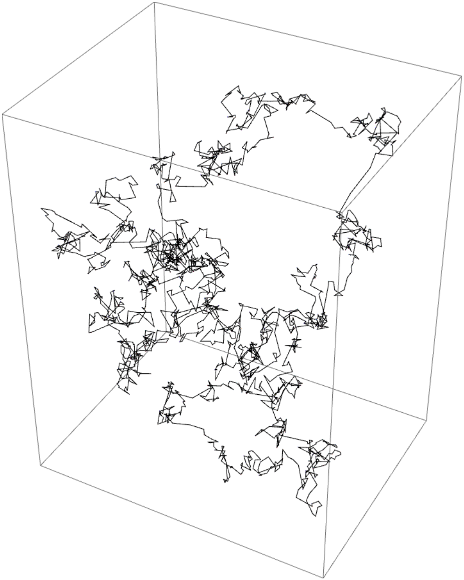
Random 2,000-gons



Random 2,000-gons



Random 2,000-gons



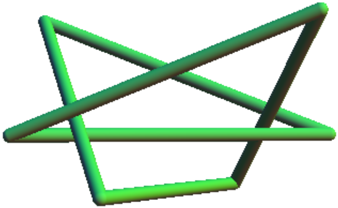
Physicists tend to model polymers with *equilateral* random walks; i.e., walks consisting of unit-length steps. The moduli space of such walks up to translation is $\underbrace{S^2(1) \times \ldots \times S^2(1)}_{n}$.

Equilateral Random Walks

Physicists tend to model polymers with *equilateral* random walks; i.e., walks consisting of unit-length steps. The moduli space of such walks up to translation is $\underbrace{S^2(1) \times \ldots \times S^2(1)}_{n}$.

Let ePol(n) be the submanifold of *closed* equilateral random walks (or *random equilateral polygons*): those walks which satisfy both $\|\vec{e}_i\| = 1$ for all i and

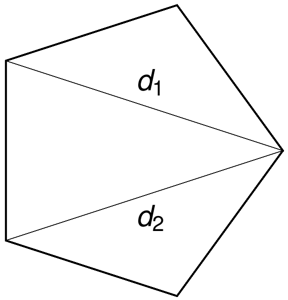
$$\sum_{i=1}^{n} \vec{e}_i = \vec{0}.$$

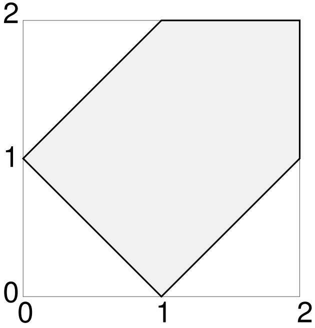


The Triangulation Polytope

Definition

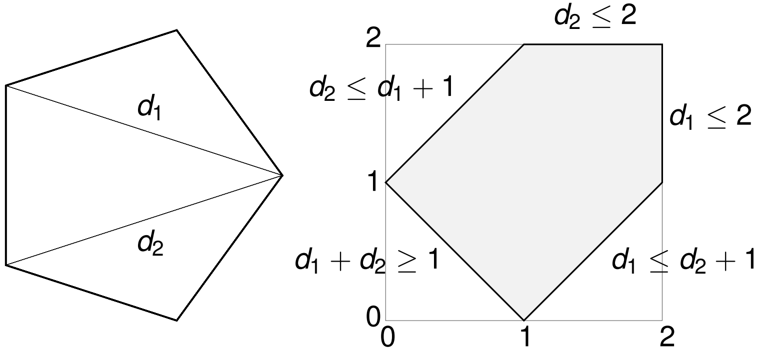
An abstract triangulation T of the n-gon picks out $n-3$ nonintersecting chords. The lengths of these chords obey triangle inequalities, so they lie in a convex polytope in $\mathbb{R}^{n-3}$ called the *triangulation polytope* $\mathcal{P}_n$.





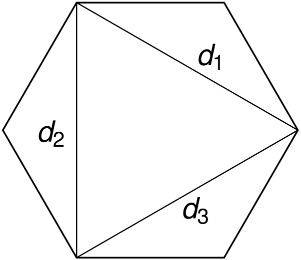
Definition

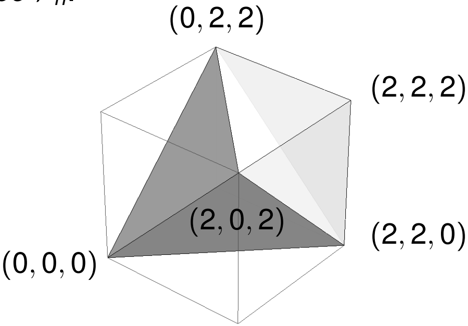
An abstract triangulation T of the n-gon picks out $n-3$ nonintersecting chords. The lengths of these chords obey triangle inequalities, so they lie in a convex polytope in $\mathbb{R}^{n-3}$ called the *triangulation polytope* $\mathcal{P}_n$.



Definition

An abstract triangulation T of the n-gon picks out $n-3$ nonintersecting chords. The lengths of these chords obey triangle inequalities, so they lie in a convex polytope in $\mathbb{R}^{n-3}$ called the *triangulation polytope* $\mathcal{P}_n$.

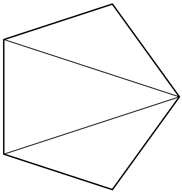


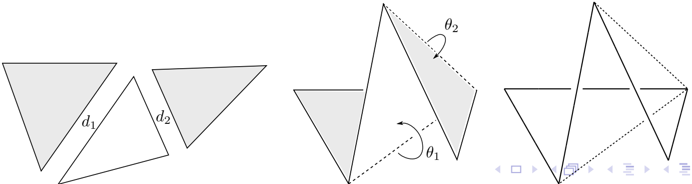


Definition

If $\mathcal{P}_n$ is the triangulation polytope and $T^{n-3} = (S^1)^{n-3}$ is the torus of $n-3$ dihedral angles, then there are *action-angle coordinates*:

$$\alpha \colon \mathcal{P}_n \times T^{n-3} \to \mathrm{ePol}(n)/\,\mathrm{SO}(3)$$





Theorem (with Cantarella)

*α pushes the **standard probability measure** on $\mathcal{P}_n \times T^{n-3}$ forward to the **correct probability measure** on $\mathrm{ePol}(n)/\,\mathrm{SO}(3)$.*

Theorem (with Cantarella)

*α pushes the **standard probability measure** on $\mathcal{P}_n \times T^{n-3}$ forward to the **correct probability measure** on $\mathrm{ePol}(n)/\,\mathrm{SO}(3)$.*

Corollary (with Cantarella)

At least $1/2$ of the space of equilateral 6-edge polygons consists of unknots.

Theorem (with Cantarella)

*α pushes the **standard probability measure** on $\mathcal{P}_n \times T^{n-3}$ forward to the **correct probability measure** on $\operatorname{ePol}(n)/\operatorname{SO}(3)$.*

Ingredients of the Proof.

Kapovich–Millson toric symplectic structure on polygon space + Duistermaat–Heckman theorem + Hitchin's theorem on compatibility of Riemannian and symplectic volume on symplectic reductions of Kähler manifolds + Howard–Manon–Millson analysis of polygon space. $\square$

Corollary (with Cantarella)

At least $1/2$ of the space of equilateral 6-edge polygons consists of unknots.

Despite the theorem, we observe experimentally that (with 95% confidence) between 1.1 and 1.5 in 10,000 hexagons are knotted.

How can we be so sure?

Despite the theorem, we observe experimentally that (with 95% confidence) between 1.1 and 1.5 in 10,000 hexagons are knotted.

How can we be so sure?

Algorithm (with Cantarella)

A Markov chain which converges to the correct measure on $\mathrm{ePol}(n)/\mathrm{SO}(3)$. *Steps in the chain are generated in* $O(n^2)$ *time. This generalizes to other fixed edgelength polygon spaces as well as to polygons in various confinement regimes.*

Despite the theorem, we observe experimentally that (with 95% confidence) between 1.1 and 1.5 in 10,000 hexagons are knotted.

How can we be so sure?

Algorithm (with Cantarella)

A Markov chain which converges to the correct measure on $\mathrm{ePol}(n)/\mathrm{SO}(3)$. *Steps in the chain are generated in* $O(n^2)$ *time. This generalizes to other fixed edgelength polygon spaces as well as to polygons in various confinement regimes.*

Algorithm (with Cantarella and Uehara)

An unbiased sampling algorithm which generates a uniform point on $\mathrm{ePol}(n)/\mathrm{SO}(3)$ *in* $O(n^{5/2})$ *time.*

Despite the theorem, we observe experimentally that (with 95% confidence) between 1.1 and 1.5 in 10,000 hexagons are knotted.

How can we be so sure?

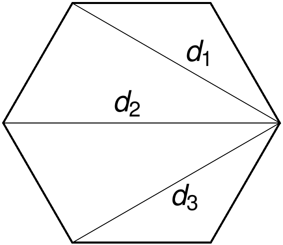
Algorithm (with Cantarella)

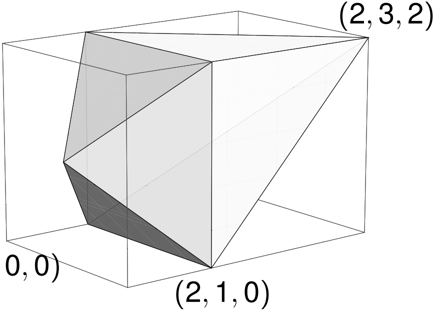
A Markov chain which converges to the correct measure on $\mathrm{ePol}(n)/\mathrm{SO}(3)$*. Steps in the chain are generated in* $O(n^2)$ *time. This generalizes to other fixed edgelength polygon spaces as well as to polygons in various confinement regimes.*

Algorithm (with Cantarella and Uehara)

An unbiased sampling algorithm which generates a uniform point on $\mathrm{ePol}(n)/\mathrm{SO}(3)$ *in* $O(n^{5/2})$ *time.*

The hard part is sampling the convex polytope $\mathcal{P}_n$.





The polytope $\mathcal{F}_n$ corresponding to the “fan triangulation” is defined by the triangle inequalities:

$$0 \le d_1 \le 2 \qquad \begin{array}{c} 1 \le d_i + d_{i+1} \\ |d_i - d_{i+1}| \le 1 \end{array} \qquad 0 \le d_{n-3} \le 2$$

A change of coordinates

If we introduce a fake chordlength $d_0 = 1$, and make the linear transformation

$$s_i = d_i - d_{i+1}, \text{ for } 0 \leq i \leq n-4, \quad s_{n-3} = d_{n-3} - d_0$$

then our inequalities

$$0 \leq d_1 \leq 2 \qquad \begin{gathered} 1 \leq d_i + d_{i+1} \\ |d_i - d_{i+1}| \leq 1 \end{gathered} \qquad 0 \leq d_{n-3} \leq 2$$

become

$$\underbrace{-1 \leq s_i \leq 1, \quad \sum s_i = 0,}_{|d_i - d_{i+1}| \leq 1} \quad \underbrace{2\sum_{j=0}^{i-1} s_j + s_i \leq 1}_{d_i + d_{i+1} \geq 1}$$

A change of coordinates

If we introduce a fake chordlength $d_0 = 1$, and make the linear transformation

$$s_i = d_i - d_{i+1}, \text{ for } 0 \leq i \leq n-4, \quad s_{n-3} = d_{n-3} - d_0$$

then our inequalities

$$0 \leq d_1 \leq 2 \qquad \begin{matrix} 1 \leq d_i + d_{i+1} \\ |d_i - d_{i+1}| \leq 1 \end{matrix} \qquad 0 \leq d_{n-3} \leq 2$$

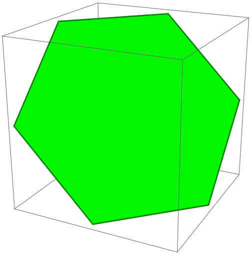
become

$$\underbrace{-1 \leq s_i \leq 1, \quad \sum s_i = 0,}_{\text{easy conditions}} \quad \underbrace{2\sum_{j=0}^{i-1} s_j + s_i \leq 1}_{\text{hard conditions}}$$

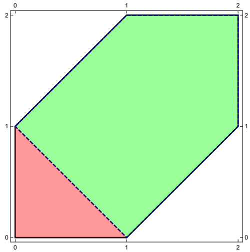
Basic Idea

Definition

The m-dimensional polytope $\mathcal{C}_m$ is the slice of the hypercube $[-1,1]^{m+1}$ by the plane $x_1 + \cdots + x_{m+1} = 0$.



$\rightarrow$



Idea

Sample points in $\mathcal{C}_{n-3}$, which all obey the "easy conditions", and reject any samples which fail to obey the "hard conditions".

Theorem (Marichal-Mossinghoff)

The volume of the projection of $\mathcal{C}_{n-3}$ is

$$\frac{\sum_{j=0}^{\lfloor \frac{n-2}{2} \rfloor} (-1)^j (n-2j-2)^{n-3} \binom{n-2}{j}}{(n-3)!}$$

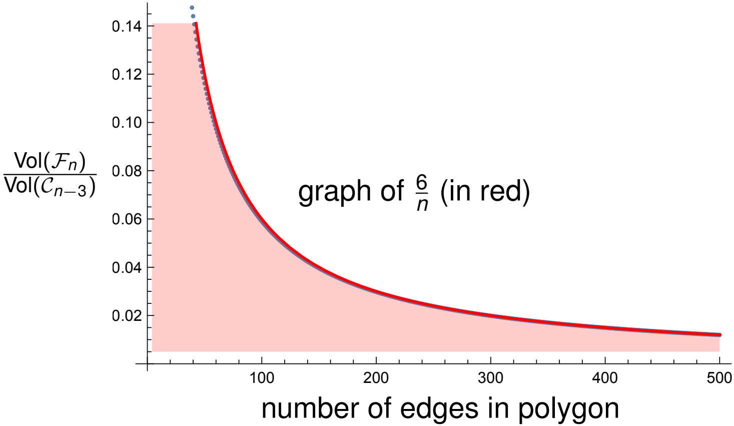
Theorem (Khoi, Takakura, Mandini)

The volume of the $(n-3)$-dimensional fan triangulation polytope for n-edge equilateral polygons $\mathcal{F}_n$ is

$$-\frac{\sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^j (n-2j)^{n-3} \binom{n}{j}}{2(n-3)!}$$

Runtime of algorithm depends on acceptance ratio

Acceptance ratio $= \frac{\mathrm{Vol}(\mathcal{F}_n)}{\mathrm{Vol}(\mathcal{C}_{n-3})}$ is conjectured $\simeq \frac{6}{n}$. It is certainly bounded below by $\frac{1}{n}$.



Recall that we're interested in the polytope $\mathcal{C}_{n-3}$ determined by

$$s_0 + s_1 + \ldots + s_{n-3} = 0$$

in the cube $[-1, 1]^{n-2}$.

Recall that we're interested in the polytope $\mathcal{C}_{n-3}$ determined by

$$s_0 + s_1 + \ldots + s_{n-3} = 0$$

in the cube $[-1, 1]^{n-2}$. After dropping the last coordinate, this corresponds to the slab

$$\mathcal{Q}_{n-3} = \{-1 \leq s_0 + s_1 + \ldots + s_{n-4} \leq 1\}$$

in the hypercube $[-1, 1]^{n-3}$.

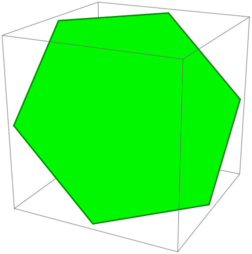
Recall that we're interested in the polytope $\mathcal{C}_{n-3}$ determined by

$$s_0 + s_1 + \ldots + s_{n-3} = 0$$

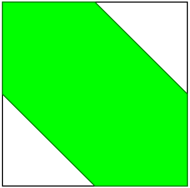
in the cube $[-1, 1]^{n-2}$. After dropping the last coordinate, this corresponds to the slab

$$\mathcal{Q}_{n-3} = \{-1 \leq s_0 + s_1 + \ldots + s_{n-4} \leq 1\}$$

in the hypercube $[-1, 1]^{n-3}$.



$\mathcal{C}_2$



$\mathcal{Q}_2$

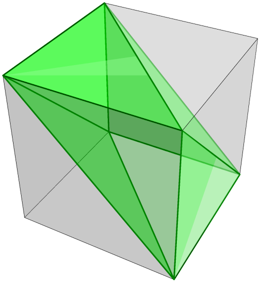
Recall that we're interested in the polytope $\mathcal{C}_{n-3}$ determined by

$$s_0 + s_1 + \ldots + s_{n-3} = 0$$

in the cube $[-1, 1]^{n-2}$. After dropping the last coordinate, this corresponds to the slab

$$\mathcal{Q}_{n-3} = \{-1 \leq s_0 + s_1 + \ldots + s_{n-4} \leq 1\}$$

in the hypercube $[-1, 1]^{n-3}$.



$\mathcal{Q}_3$

We can sample $\mathcal{Q}_{n-3}$ (and hence $\mathcal{C}_{n-3}$) by rejection sampling $[-1,1]^{n-3}$. The acceptance probability is

$$\frac{\operatorname{Vol} \mathcal{Q}_{n-3}}{\operatorname{Vol}[-1,1]^{n-3}}.$$

We can sample $\mathcal{Q}_{n-3}$ (and hence $\mathcal{C}_{n-3}$) by rejection sampling $[-1, 1]^{n-3}$. The acceptance probability is

$$\frac{\text{Vol}\,\mathcal{Q}_{n-3}}{\text{Vol}[-1, 1]^{n-3}}.$$

But this is just the probability that a sum of Uniform($[-1, 1]$) variates is between -1 and 1.

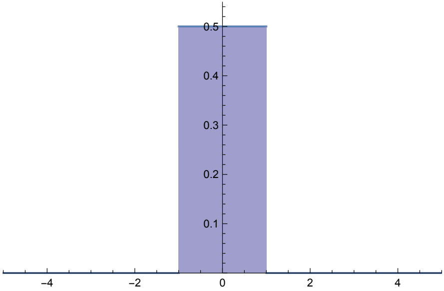
We can sample $\mathcal{Q}_{n-3}$ (and hence $\mathcal{C}_{n-3}$) by rejection sampling $[-1,1]^{n-3}$. The acceptance probability is

$$\frac{\operatorname{Vol} \mathcal{Q}_{n-3}}{\operatorname{Vol}[-1,1]^{n-3}}.$$

But this is just the probability that a sum of Uniform($[-1,1]$) variates is between -1 and 1.

$n = 1$

$\mathbb{P}(-1 \sum s_i \leq 1) = 1$



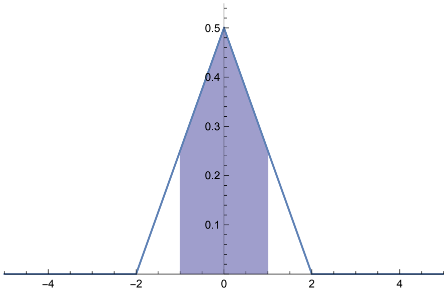
We can sample $\mathcal{Q}_{n-3}$ (and hence $\mathcal{C}_{n-3}$) by rejection sampling $[-1,1]^{n-3}$. The acceptance probability is

$$\frac{\text{Vol}\,\mathcal{Q}_{n-3}}{\text{Vol}[-1,1]^{n-3}}.$$

But this is just the probability that a sum of Uniform($[-1,1]$) variates is between -1 and 1.

$n = 2$

$\mathbb{P}(-1 \sum s_i \leq 1) = \frac{3}{4}$



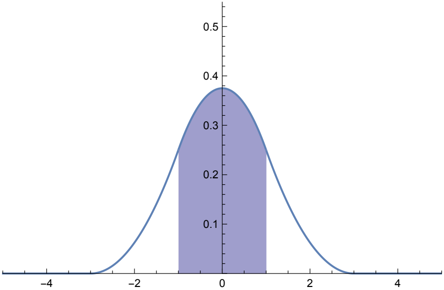
We can sample $\mathcal{Q}_{n-3}$ (and hence $\mathcal{C}_{n-3}$) by rejection sampling $[-1,1]^{n-3}$. The acceptance probability is

$$\frac{\mathrm{Vol}\,\mathcal{Q}_{n-3}}{\mathrm{Vol}[-1,1]^{n-3}}.$$

But this is just the probability that a sum of Uniform($[-1,1]$) variates is between -1 and 1.

$n = 3$

$\mathbb{P}(-1 \sum s_i \leq 1) = \frac{2}{3}$



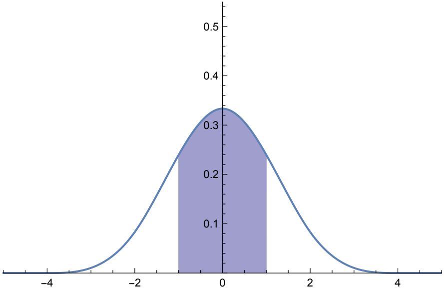
We can sample $\mathcal{Q}_{n-3}$ (and hence $\mathcal{C}_{n-3}$) by rejection sampling $[-1,1]^{n-3}$. The acceptance probability is

$$\frac{\text{Vol}\,\mathcal{Q}_{n-3}}{\text{Vol}[-1,1]^{n-3}}.$$

But this is just the probability that a sum of Uniform($[-1,1]$) variates is between -1 and 1.

$n = 4$

$\mathbb{P}(-1 \sum s_i \leq 1) = \frac{115}{192}$



Sampling $\mathcal{Q}_{n-3}$

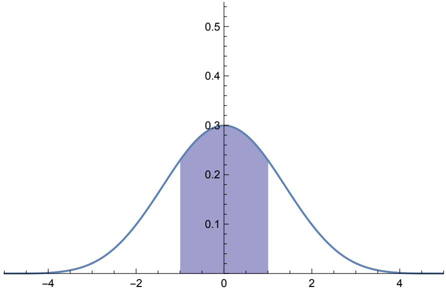
We can sample $\mathcal{Q}_{n-3}$ (and hence $\mathcal{C}_{n-3}$) by rejection sampling $[-1,1]^{n-3}$. The acceptance probability is

$$\frac{\mathrm{Vol}\,\mathcal{Q}_{n-3}}{\mathrm{Vol}[-1,1]^{n-3}}.$$

But this is just the probability that a sum of Uniform($[-1,1]$) variates is between -1 and 1.

$n = 5$

$\mathbb{P}(-1 \sum s_i \leq 1) = \frac{11}{20}$



Sampling $\mathcal{Q}_{n-3}$

We can sample $\mathcal{Q}_{n-3}$ (and hence $\mathcal{C}_{n-3}$) by rejection sampling $[-1,1]^{n-3}$. The acceptance probability is

$$\frac{\text{Vol}\,\mathcal{Q}_{n-3}}{\text{Vol}[-1,1]^{n-3}}.$$

But this is just the probability that a sum of Uniform($[-1,1]$) variates is between -1 and 1.

By Shepp's **local limit theorem**,

$$\mathbb{P}\left(-1 \le \sum_{i=0}^{n-4} s_i \le 1\right) \simeq \mathbb{P}\left(-1 \le \mathcal{N}\left(0, \sqrt{\frac{n-3}{3}}\right) \le 1\right)$$

We can sample $\mathcal{Q}_{n-3}$ (and hence $\mathcal{C}_{n-3}$) by rejection sampling $[-1,1]^{n-3}$. The acceptance probability is

$$\frac{\mathrm{Vol}\,\mathcal{Q}_{n-3}}{\mathrm{Vol}[-1,1]^{n-3}}.$$

But this is just the probability that a sum of Uniform($[-1,1]$) variates is between -1 and 1.

By Shepp's **local limit theorem**,

$$\mathbb{P}\left(-1 \le \sum_{i=0}^{n-4} s_i \le 1\right) \simeq \mathbb{P}\left(-1 \le \mathcal{N}\left(0, \sqrt{\frac{n-3}{3}}\right) \le 1\right)$$

$$= \mathrm{erf}\left(\sqrt{\frac{3}{2(n-3)}}\right) \simeq \sqrt{\frac{6}{\pi n}}.$$

Moment Polytope Sampling Algorithm (with Cantarella and Uehara, 2015)

1. *Generate* $(s_0, \ldots, s_{n-4})$ *uniformly on* $[-1, 1]^{n-3}$ $O(n)$ *time*
2. *Test whether* $-1 \leq \sum s_i \leq 1$ *acceptance ratio* $\simeq 1/\sqrt{n}$
3. *Let* $s_{n-3} = -\sum s_i$ *and test* $(s_0, \ldots, s_{n-3})$ *against the "hard" conditions* *acceptance ratio* $> 1/n$
4. *Change coordinates to get diagonal lengths*
5. *Generate dihedral angles from* T^{n-3}
6. *Build sample polygon in action-angle coordinates*

Future Challenges

- Topologically-constrained random walks based on more complicated graphs.
- Incorporate geometric constraints (restricted turning angles, excluded volume, etc.), which correspond to physical restrictions on polymer geometries.
- Stronger theoretical bounds on knot probabilities, average knot invariants, etc.
- A general theory of random piecewise-linear submanifolds?

Thank you!

Thank you for listening!

References

- *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

- *The Symplectic Geometry of Closed Equilateral Random Walks in 3-Space*
  Jason Cantarella and Clayton Shonkwiler
  Annals of Applied Probability, to appear.

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

Recall the volumes of the cross polytope and moment polytope:

$$\mathrm{Vol}(\mathcal{C}_{n-3}) = \frac{\sum_{j=0}^{\left\lfloor \frac{n-2}{2} \right\rfloor} (-1)^j (n-2j-2)^{n-3} \binom{n-2}{j}}{(n-3)!}$$

$$\mathrm{Vol}(\mathcal{F}_n) = -\frac{\sum_{j=0}^{\left\lfloor \frac{n}{2} \right\rfloor} (-1)^j (n-2j)^{n-3} \binom{n}{j}}{2(n-3)!}$$

Recall the volumes of the cross polytope and moment polytope:

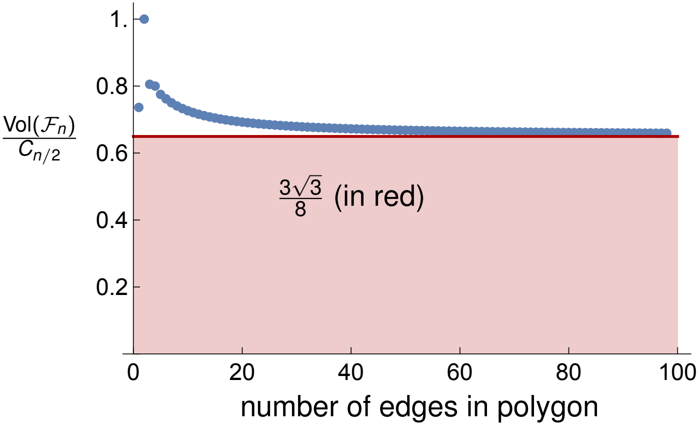
$$\mathrm{Vol}(\mathcal{F}_n) = -\frac{\sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^j (n-2j)^{n-3} \binom{n}{j}}{2(n-3)!}$$

$$\underset{\text{conj.}}{\simeq} \frac{3\sqrt{3}}{8} C_{n/2}$$

where C_m is the mth Catalan number

$$C_m = \frac{1}{m+1}\binom{2m}{m} = \frac{\Gamma(2m+1)}{\Gamma(m+2)\Gamma(m+1)}.$$

The ratio of $\mathrm{Vol}(\mathcal{F}_n)$ to $C_{n/2}$:



A Recurrence Relation for $\mathrm{Vol}(\mathcal{F}_n)$

The normalized volume $(n-3)!\,\mathrm{Vol}(\mathcal{F}_n)$ of the moment polytope is the $k = 1$ case of a two-parameter family $V(n, k)$ given by the recurrence relation

$$V(n,k) = (n-k-1)\,V(n-1,k-1) + (n+k-1)\,V(n-1,k+1)$$

subject to the boundary conditions

$$V(n,0) = 0, \qquad V(3,1) = 1, \qquad V(3,2) = \frac{1}{2}.$$

The normalized volume $(n-3)!\,\mathrm{Vol}(\mathcal{F}_n)$ of the moment polytope is the $k=1$ case of a two-parameter family $V(n,k)$ given by the recurrence relation

$$V(n,k) = (n-k-1)\,V(n-1,k-1) + (n+k-1)\,V(n-1,k+1)$$

0	1		$\frac{1}{2}$
0			
0			
0			
0			
0			
0			
0			

A Recurrence Relation for $\mathrm{Vol}(\mathcal{F}_n)$

The normalized volume $(n-3)!\,\mathrm{Vol}(\mathcal{F}_n)$ of the moment polytope is the $k=1$ case of a two-parameter family $V(n,k)$ given by the recurrence relation

$$V(n,k) = (n-k-1)V(n-1,k-1) + (n+k-1)V(n-1,k+1)$$

0	1	$\frac{1}{2}$						
0	2	1						
0	5	4	1					
0	24	22	8	1				
0	154	160	75	16	1			
0	1280	1445	800	236	32	1		
0	13005	15680	9821	3584	721	64	1	
0	156800	199066	137088	58478	15232	2178	128	1

cf. OEIS A012249 “Volume of a certain rational polytope. . . ”

Conjecture

$$\mathrm{Vol}(\mathcal{F}_n) \simeq \frac{3\sqrt{3}}{8} C_{n/2}.$$

Conjecture

$$\mathrm{Vol}(\mathcal{F}_n) \simeq \frac{3\sqrt{3}}{8} C_{n/2}.$$

Combining the conjecture with Stirling's approximation and

$$\mathrm{Vol}(\mathcal{C}_{n-3}) \simeq 2^{n-3}\sqrt{\frac{6}{\pi n}}$$

would suffice to prove

$$\frac{\mathrm{Vol}(\mathcal{F}_n)}{\mathrm{Vol}(\mathcal{C}_{n-3})} \simeq \frac{6}{n}.$$