Closed Random Walks and Symplectic Geometry

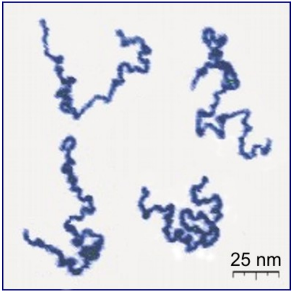
Clayton Shonkwiler

University of Georgia

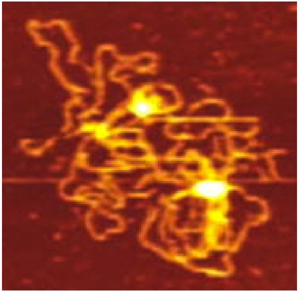
University of Rochester
January 28, 2014

Physics Question

What is the average shape of a polymer in solution?



Protonated P2VP
Roiter/Minko
Clarkson University



Plasmid DNA
Alonso-Sarduy, Dietler Lab
EPF Lausanne

Physics Question

What is the average shape of a polymer in solution?

Physics Answer

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

Let $\mathrm{Arm}(n)$ be the moduli space of random walks in $\mathbb{R}^3$ consisting of n unit-length steps up to translation.

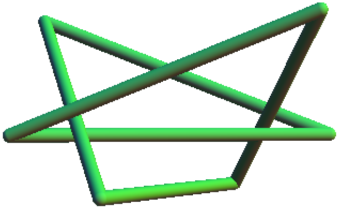
Then $\mathrm{Arm}(n) \cong \underbrace{S^2(1) \times \ldots \times S^2(1)}_{n}$.

Let Arm(n) be the moduli space of random walks in $\mathbb{R}^3$ consisting of n unit-length steps up to translation.

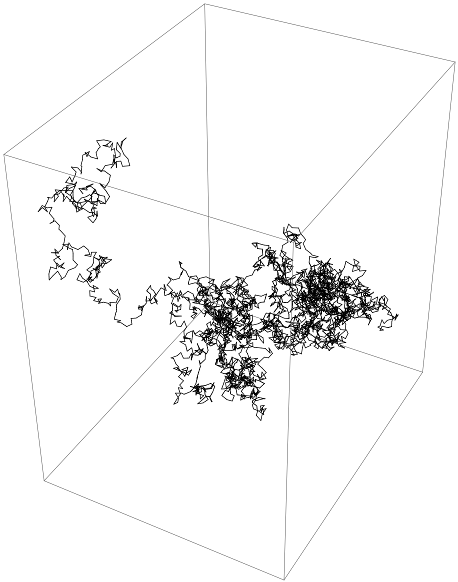
Then $\mathrm{Arm}(n) \cong \underbrace{S^2(1) \times \ldots \times S^2(1)}_{n}$.

Let Pol(n) $\subset$ Arm(n) be the submanifold of closed random walks (or *random polygons*); i.e., those walks which satisfy

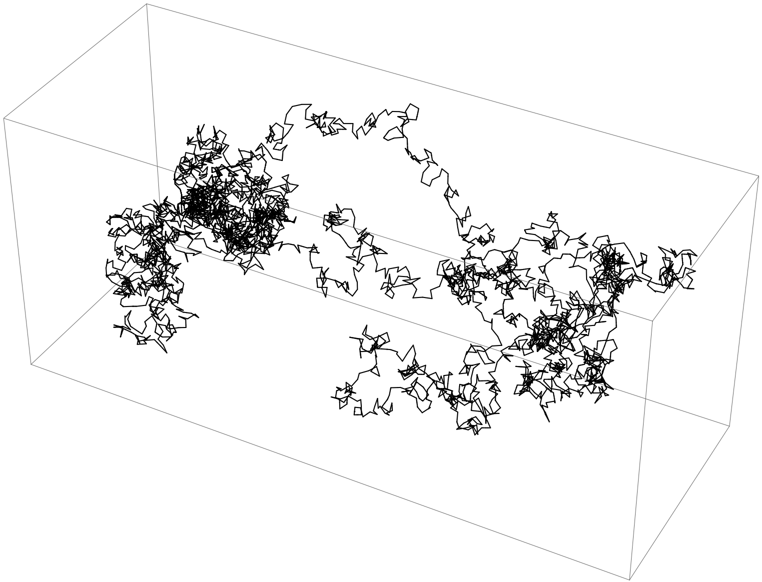
$$\sum_{i=1}^{n} \vec{e}_i = \vec{0}.$$



A Random Walk with 3,500 Steps



A Closed Random Walk with 3,500 Steps



- What is the joint distribution of steps in a closed walk?

- What is the joint distribution of steps in a closed walk?
- What can we prove about closed random walks?
    - What is the marginal distribution of a single chord length?
    - What is the joint distribution of several chord lengths?
    - What is the expectation of radius of gyration?
    - What is the expectation of total curvature?

Classical Problems

- What is the joint distribution of steps in a closed walk?
- What can we prove about closed random walks?
    - What is the marginal distribution of a single chord length?
    - What is the joint distribution of several chord lengths?
    - What is the expectation of radius of gyration?
    - What is the expectation of total curvature?
- How do we sample closed equilateral random walks?
    - What if the walk is confined to a sphere? (Confined DNA)
    - Can we get error bars?

- What is the joint distribution of steps in a closed walk?
- What can we prove about closed random walks?
  - What is the marginal distribution of a single chord length?
  - What is the joint distribution of several chord lengths?
  - What is the expectation of radius of gyration?
  - What is the expectation of total curvature?
- How do we sample closed equilateral random walks?
  - What if the walk is confined to a sphere? (Confined DNA)
  - Can we get error bars?

Point of Talk

New sampling algorithms backed by deep and robust mathematical framework. Guaranteed to converge, relatively easy to code.

(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)

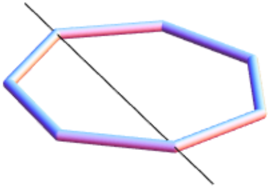
(Incomplete?) History of Sampling Algorithms

- 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

Definition

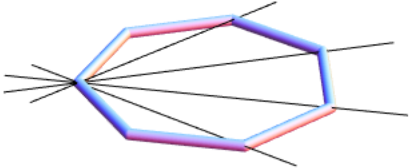
A *fold move* or *bending flow* rotates an arc of the polygon around the axis determined by its endpoints.

The polygonal fold Markov chain selects arcs and angles at random and folds repeatedly.



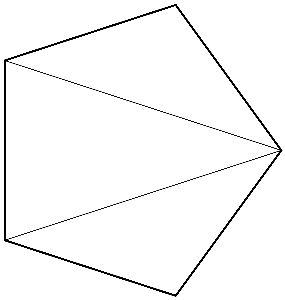
Definition

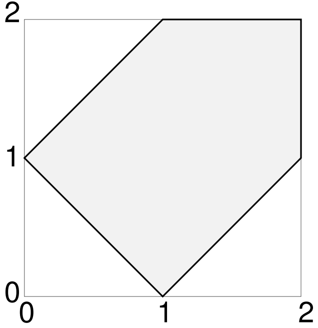
Given an (abstract) triangulation of the n-gon, the folds on any two chords commute. A *dihedral angle move* rotates around all of these chords by independently selected angles.



Definition

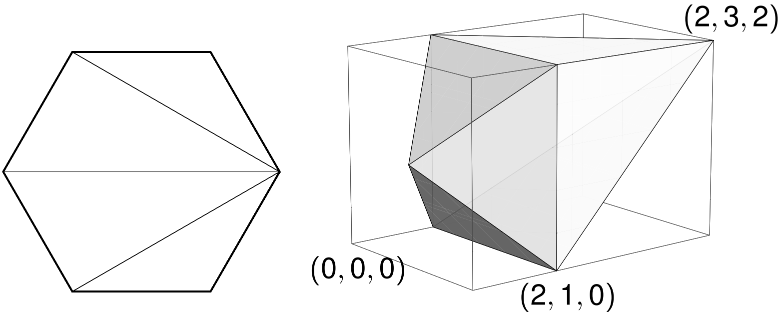
A abstract triangulation T of the n-gon picks out $n-3$ nonintersecting chords. The lengths of these chords obey triangle inequalities, so they lie in a convex polytope in $\mathbb{R}^{n-3}$ called the *triangulation polytope* $\mathcal{P}$.





Definition

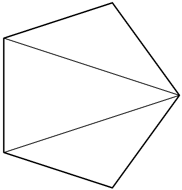
A abstract triangulation T of the n-gon picks out $n-3$ nonintersecting chords. The lengths of these chords obey triangle inequalities, so they lie in a convex polytope in $\mathbb{R}^{n-3}$ called the *triangulation polytope* $\mathcal{P}$.

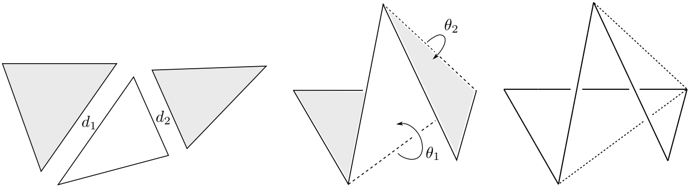


Definition

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

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





Theorem (with Cantarella)

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

Theorem (with Cantarella)

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

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. $\square$

Theorem (Archimedes, Duistermaat–Heckman)

Let $f : S^2 \to \mathbb{R}$ be given by $(x, y, z) \mapsto z$. Then the pushforward of the standard measure on the sphere to the interval is 2π times Lebesgue measure.

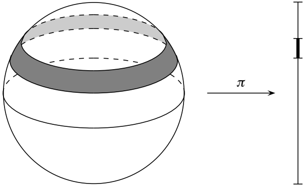
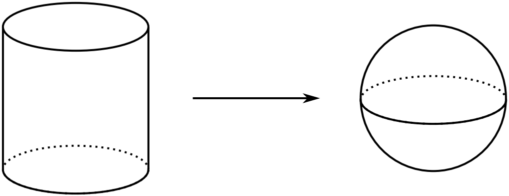


Illustration by Kuperberg.

Action-Angle Coordinates are Cylindrical Coordinates

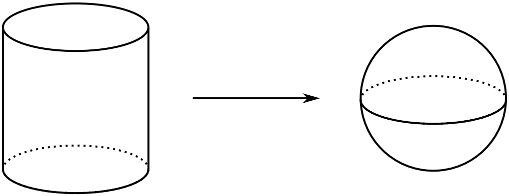


(z, θ) $\qquad$ $(\sqrt{1-z^2}\cos\theta, \sqrt{1-z^2}\sin\theta, z)$

Corollary

This map pushes the standard probability measure on $[-1, 1] \times S^1$ forward to the correct probability measure on S^2.

Action-Angle Coordinates are Cylindrical Coordinates



(z, θ) $\quad$ $(\sqrt{1-z^2}\cos\theta, \sqrt{1-z^2}\sin\theta, z)$

Corollary

More generally, the standard probability measure on $[-1,1]^n \times T^n$ pushes forward to the correct probability measure on $(S^2)^n = \mathrm{Arm}(n)$.

Main Theorem Repeated

Theorem (with Cantarella)

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

Corollary

Any sampling algorithm for convex polytopes is a sampling algorithm for closed equilateral polygons.

Corollary

The expectation of any function of a collection of non-intersecting chordlengths can be computed by integrating over the triangulation polytope.

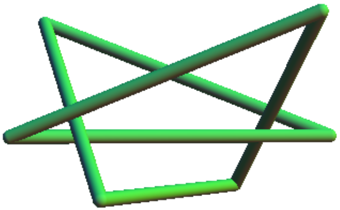
Theorem (with Cantarella)

The expected length of a chord skipping k edges in an n-edge closed equilateral random walk is the $(k-1)$st coordinate of the center of mass of the triangulation polytope for $\mathrm{Pol}(n)$.

$n \backslash k$	2	3	4	5	6	7	8
4	1						
5	$\frac{17}{15}$	$\frac{17}{15}$					
6	$\frac{14}{12}$	$\frac{15}{12}$	$\frac{14}{12}$				
7	$\frac{461}{385}$	$\frac{506}{385}$	$\frac{506}{385}$	$\frac{461}{385}$			
8	$\frac{1{,}168}{960}$	$\frac{1{,}307}{960}$	$\frac{1{,}344}{960}$	$\frac{1{,}307}{960}$	$\frac{1{,}168}{960}$		
9	$\frac{112{,}121}{91{,}035}$	$\frac{127{,}059}{91{,}035}$	$\frac{133{,}337}{91{,}035}$	$\frac{133{,}337}{91{,}035}$	$\frac{127{,}059}{91{,}035}$	$\frac{112{,}121}{91{,}035}$	
10	$\frac{97{,}456}{78{,}400}$	$\frac{111{,}499}{78{,}400}$	$\frac{118{,}608}{78{,}400}$	$\frac{120{,}985}{78{,}400}$	$\frac{118{,}608}{78{,}400}$	$\frac{111{,}499}{78{,}400}$	$\frac{97{,}456}{78{,}400}$

Theorem (with Cantarella)

At least $1/2$ of six-step closed random walks are unknotted.



Proof.

Using a result of Calvo, the possible dihedral angles of a knotted hexagon comprise no more than half of the torus of dihedrals. □

Recall

Action-angle coordinates reduce sampling closed random walks to the (solved) problem of sampling a convex polytope.

Recall

Action-angle coordinates reduce sampling closed random walks to the (solved) problem of sampling a convex polytope.

Definition (Hit-and-run Sampling Markov Chain)

Given $\vec{p}_k \in \mathcal{P} \subset \mathbb{R}^n$,

1. Choose a random direction $\vec{v}$ uniformly on S^{n-1}.
2. Let ℓ be the line through $\vec{p}_k$ in direction $\vec{v}$.
3. Choose $\vec{p}_{k+1}$ uniformly on $\ell \cap \mathcal{P}$.

A Markov Chain for Convex Polytopes

Recall

Action-angle coordinates reduce sampling closed random walks to the (solved) problem of sampling a convex polytope.

Definition (Hit-and-run Sampling Markov Chain)

Given $\vec{p}_k \in \mathcal{P} \subset \mathbb{R}^n$,

1. Choose a random direction $\vec{v}$ uniformly on S^{n-1}.
2. Let ℓ be the line through $\vec{p}_k$ in direction $\vec{v}$.
3. Choose $\vec{p}_{k+1}$ uniformly on $\ell \cap \mathcal{P}$.

Theorem (Smith, 1984)

The hit-and-run Markov chain is uniformly ergodic with respect to Lebesgue measure on $\mathcal{P}$.

A (new) Markov Chain for Closed Random Walks

Definition (TSMCMC(β))

If $x_k = (\vec{p}_k, \vec{\theta}_k) \in \mathcal{P} \times T^{n-3}$, define x_{k+1} by:

- With probability β, update $\vec{p}_k$ by a hit-and-run step on $\mathcal{P}$.
- With probability $1 - \beta$, replace $\vec{\theta}_k$ with a new uniformly sampled point in T^{n-3}.

At each step, construct the corresponding closed random walk $\alpha(x_k)$ using action-angle coordinates.

A (new) Markov Chain for Closed Random Walks

Definition (TSMCMC(β))

If $x_k = (\vec{p}_k, \vec{\theta}_k) \in \mathcal{P} \times T^{n-3}$, define x_{k+1} by:

- With probability β, update $\vec{p}_k$ by a hit-and-run step on $\mathcal{P}$.
- With probability $1 - \beta$, replace $\vec{\theta}_k$ with a new uniformly sampled point in T^{n-3}.

At each step, construct the corresponding closed random walk $\alpha(x_k)$ using action-angle coordinates.

Proposition (with Cantarella)

TSMCMC(β) is uniformly ergodic with respect to the standard probability measure on $\mathrm{Pol}(n)/SO(3)$.

Error Analysis for Integration with TSMCMC(β)

Suppose f is a function on closed random walks and SampleMean($f; R, m$) is the average value of f over the first m steps of a run R of *TSMCMC*(β).

Because TSMCMC(β) is uniformly ergodic, we have

[1] w denotes weak convergence, $E(f)$ is the expectation of f

Error Analysis for Integration with TSMCMC(β)

Suppose f is a function on closed random walks and SampleMean$(f; R, m)$ is the average value of f over the first m steps of a run R of *TSMCMC*(β).

Because TSMCMC(β) is uniformly ergodic, we have

Theorem (Markov Chain Central Limit Theorem)

If f is square-integrable, there exists a real number $\sigma(f)$ so that[1]

$$\sqrt{m}(\text{SampleMean}(f; R, m) - E(f)) \xrightarrow{w} \mathcal{N}(0, \sigma(f)^2),$$

the Gaussian with mean 0 and standard deviation $\sigma(f)^2$.

[1] w denotes weak convergence, $E(f)$ is the expectation of f

Error Bars for TSMCMC(β) Integration

1. Run TSMCMC(β) for m steps
2. Compute SampleMean($f; R, m$)
3. Compute the **Geyer IPS Estimator** $\bar{\sigma}_m(f)$ for $\sigma(f)$

Then with 95% confidence the true expectation of f is

$$E(f) = \text{SampleMean}(f; R, m) \pm 1.96\bar{\sigma}_m(f)/\sqrt{m}.$$

1. Run TSMCMC(β) for m steps
2. Compute SampleMean(f; R, m)
3. Compute the **Geyer IPS Estimator** $\bar{\sigma}_m(f)$ for $\sigma(f)$

Then with 95% confidence the true expectation of f is

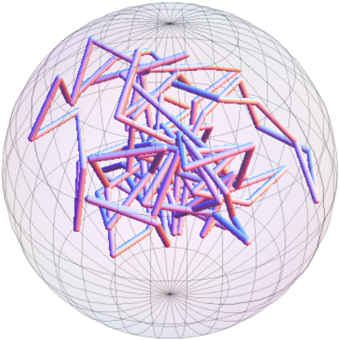
$$E(f) = \text{SampleMean}(f; R, m) \pm 1.96\bar{\sigma}_m(f)/\sqrt{m}.$$

Experimental Observation

With 95% confidence, we can say that the fraction of knotted equilateral hexagons is between 1.1 and 1.5 in 10,000.

Definition

A closed random walk is in *rooted spherical confinement* of radius R if each chord length is less than R. Such a polygon is contained in a sphere of radius R centered at the first vertex.



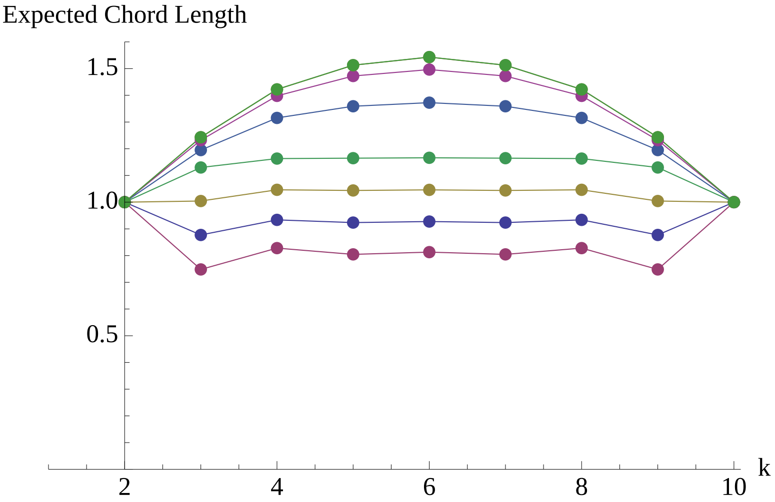
Let $\mathcal{P}_R$ be the subpolytope of $\mathcal{P}$ determined by the additional linear conditions that each chord length is less than R.

Theorem (with Cantarella)

*The action-angle map α pushes the **standard probability measure** on $\mathcal{P}_R \times T^{n-3}$ forward to the **correct probability measure** on the space of closed random walks in rooted spherical confinement of radius R.*

Many other confinement models are possible!

Expected Chordlengths for Confined 10-gons



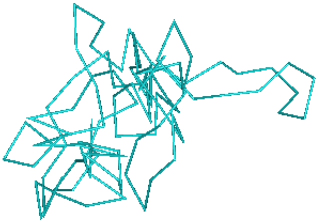
Confinement radii are 1.25, 1.5, 1.75, 2, 2.5, 3, 4, and 5.

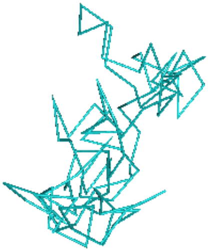
Unconfined 100-gons

20-confined 100-gons

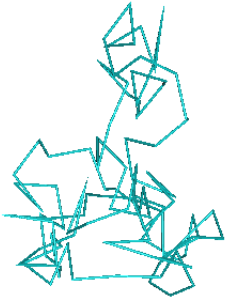
20-confined 100-gons

10-confined 100-gons





10-confined 100-gons



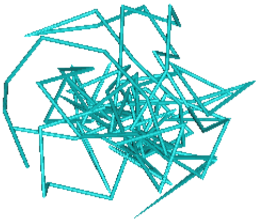
5-confined 100-gons

5-confined 100-gons

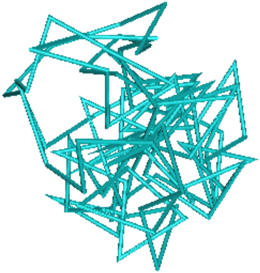
5-confined 100-gons

5-confined 100-gons

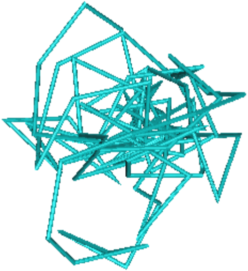
2-confined 100-gons

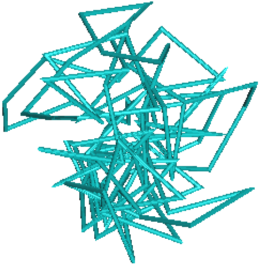


2-confined 100-gons

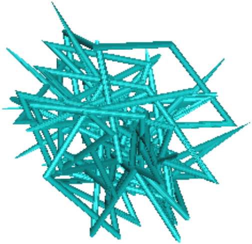


2-confined 100-gons

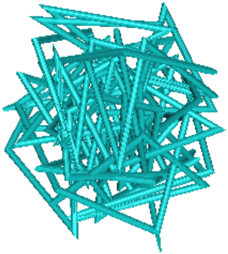




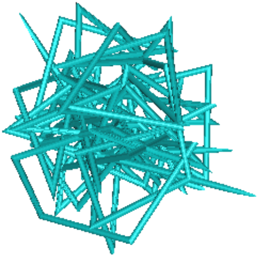
1.1-confined 100-gons

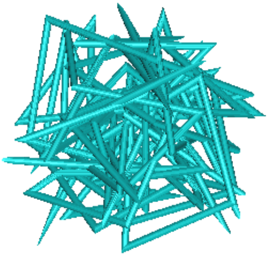


1.1-confined 100-gons



1.1-confined 100-gons





Thank you for listening!

References

- *Probability Theory of Random Polygons from the Quaternionic Viewpoint*
  Jason Cantarella, Tetsuo Deguchi, and Clayton Shonkwiler
  arXiv:1206.3161
  Communications on Pure and Applied Mathematics
  (2013), doi:10.1002/cpa.21480.

- *The Expected Total Curvature of Random Polygons*
  Jason Cantarella, Alexander Y. Grosberg, Robert Kusner, and Clayton Shonkwiler
  arXiv:1210.6537.

- *The symplectic geometry of closed equilateral random walks in 3-space*
  Jason Cantarella and Clayton Shonkwiler
  arXiv:1310.5924.