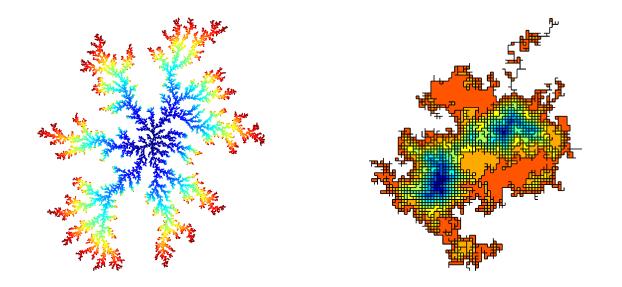
# RANDOM THOUGHTS ON RANDOM SETS Christopher Bishop, Stony Brook

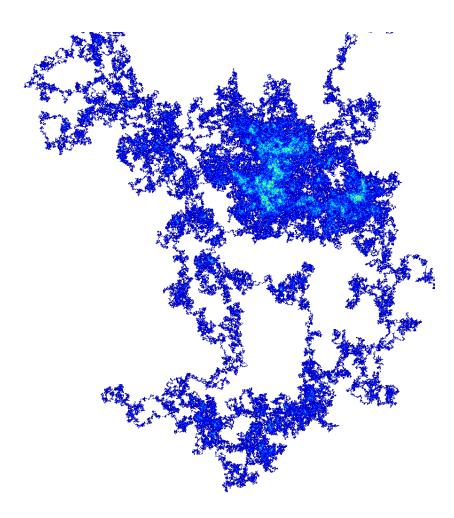
#### April 7, 2021

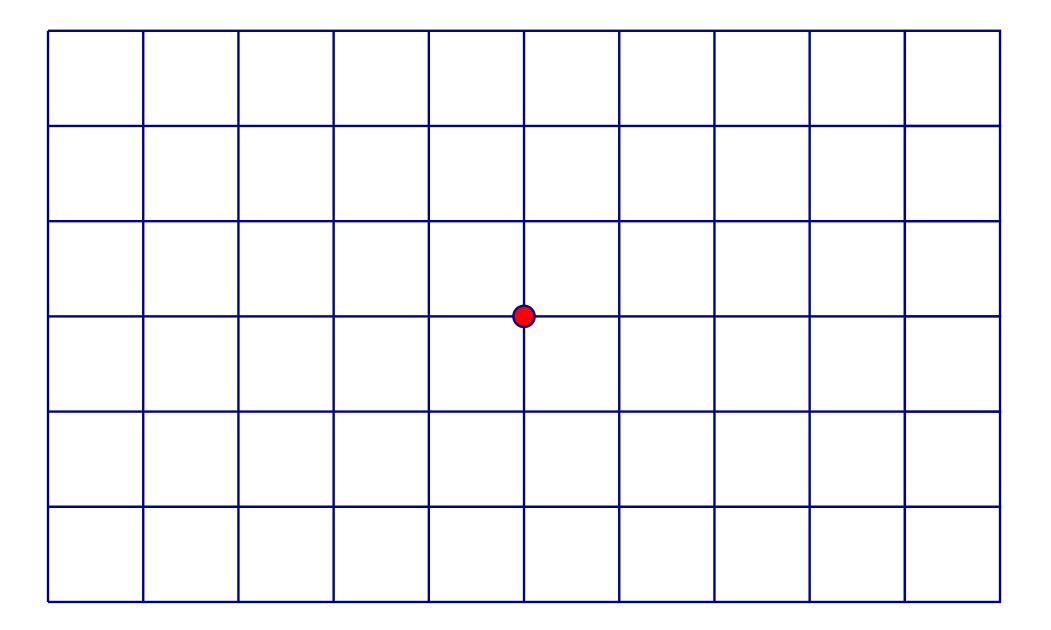
## Stony Brook Math Club and Indiana University Bloomington Math Club

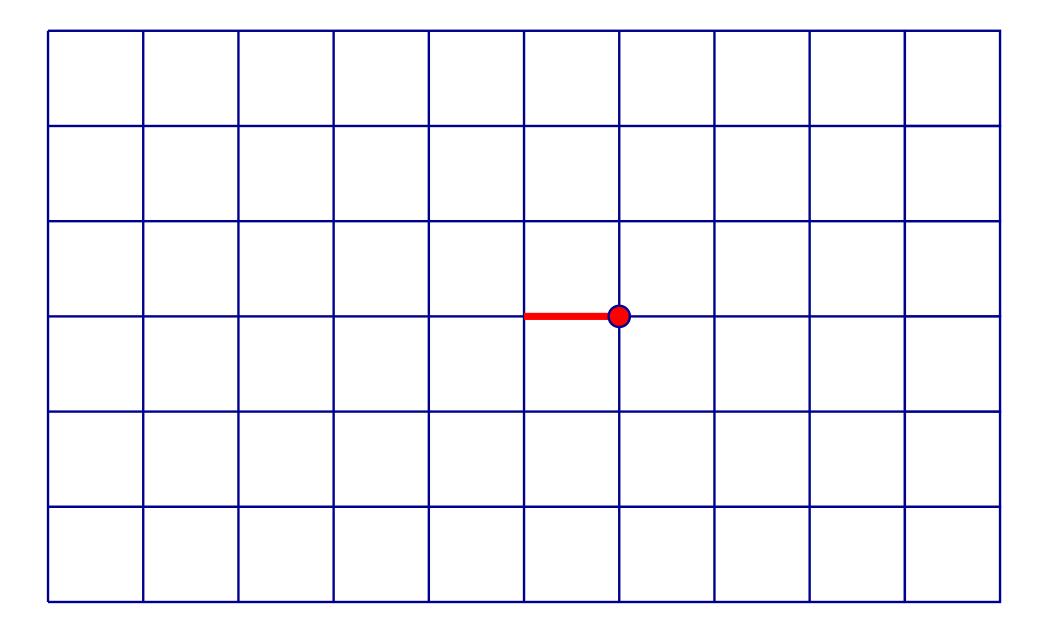
www.math.sunysb.edu/~bishop/lectures

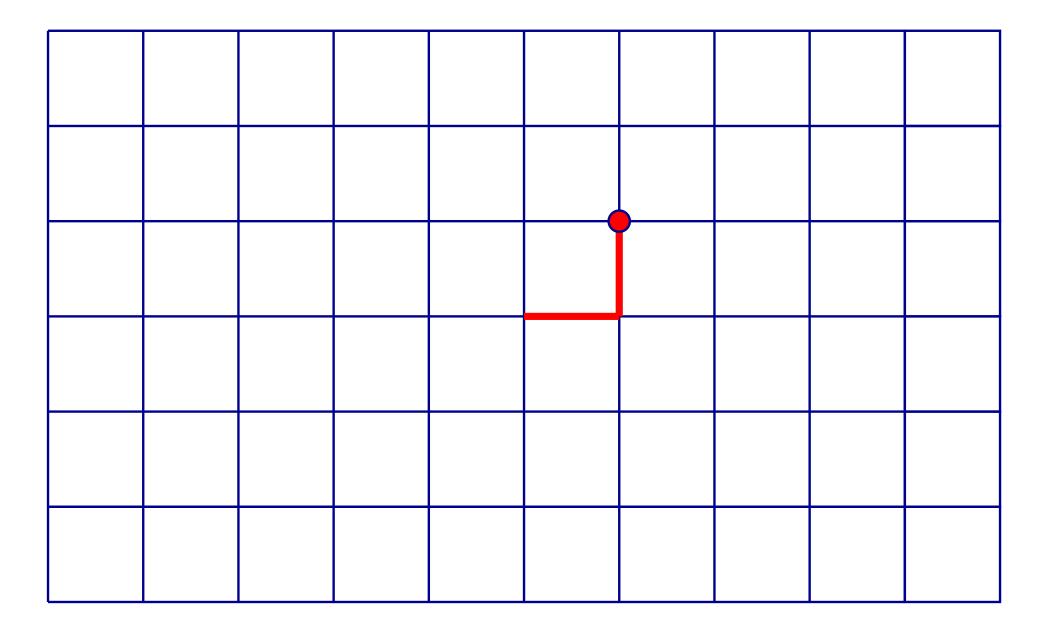


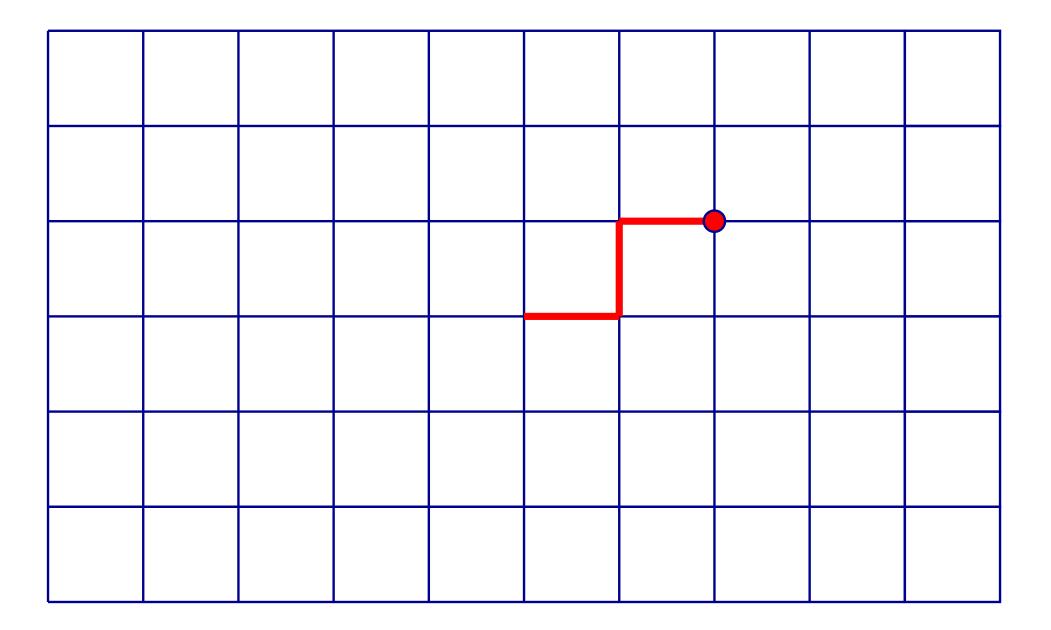
#### PART I: RANDOM WALKS AND BROWNIAN MOTION

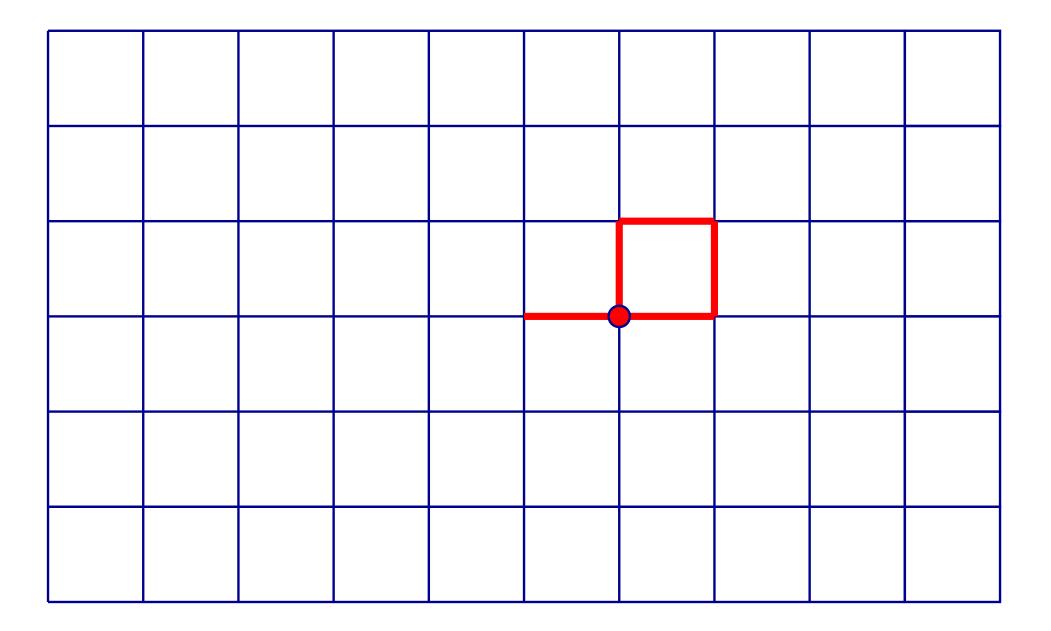


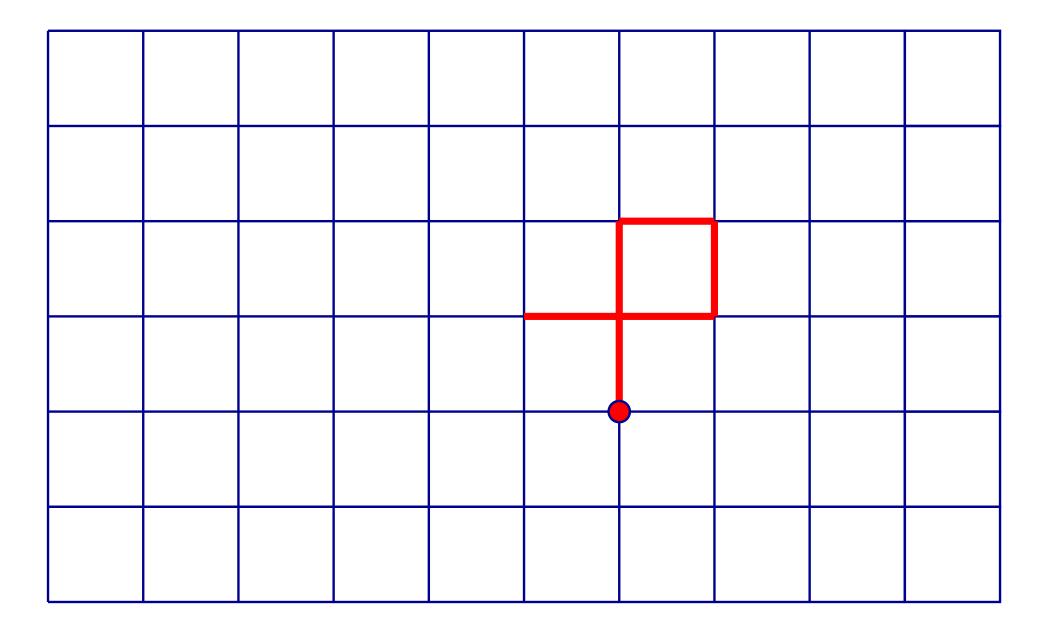


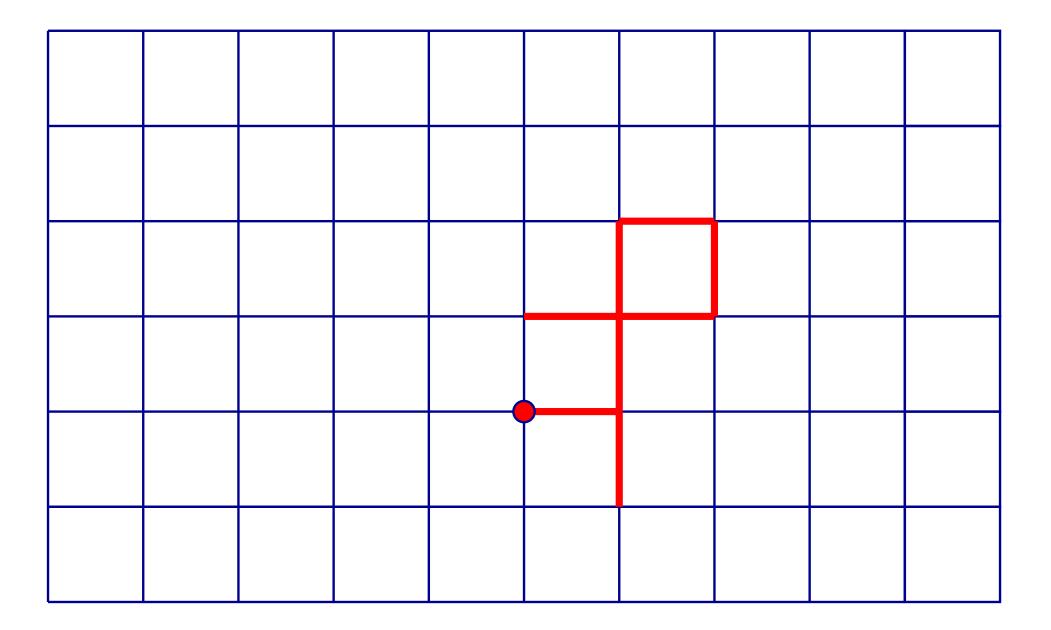


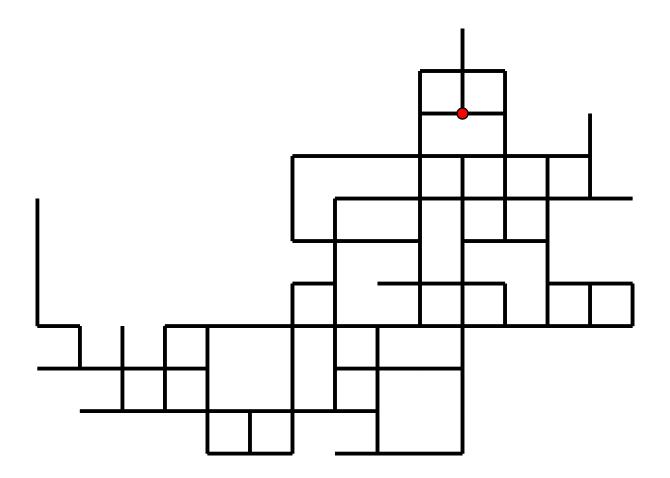




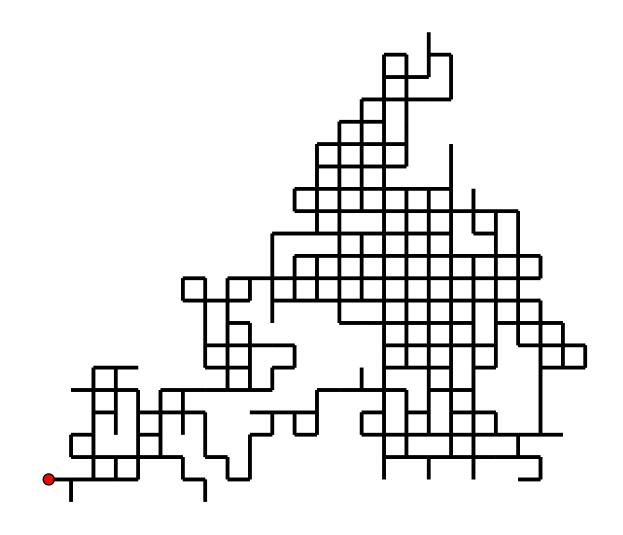




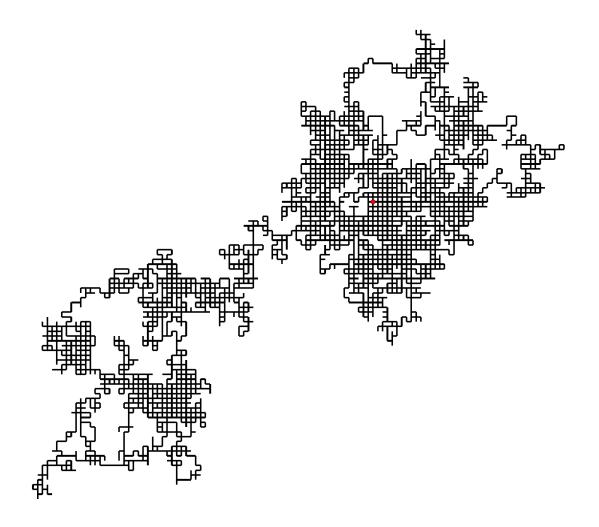




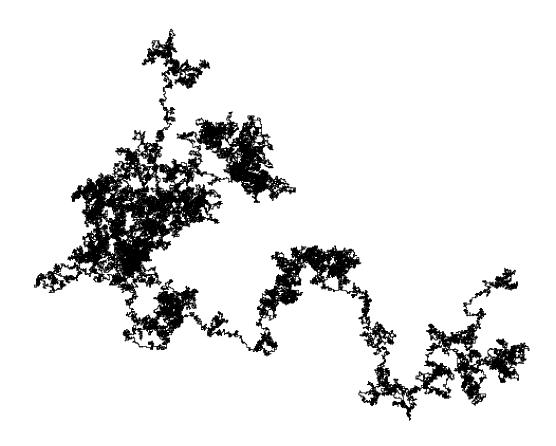
200 step random walk.



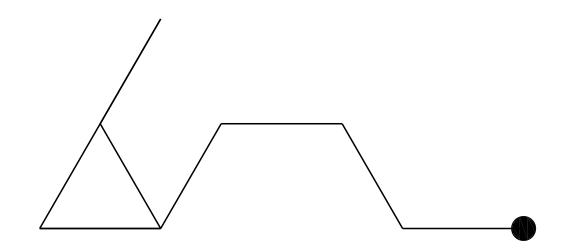
1000 step random walk.



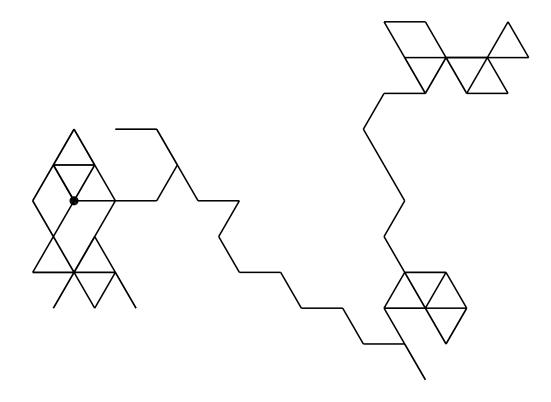
10,000 step random walk.



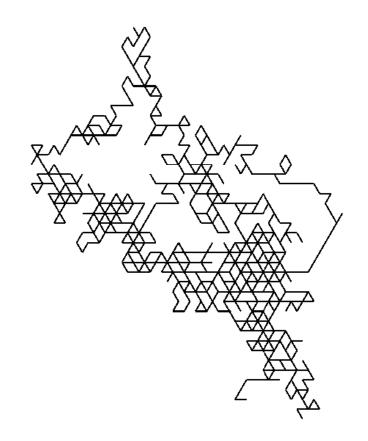
100,000 step random walk.



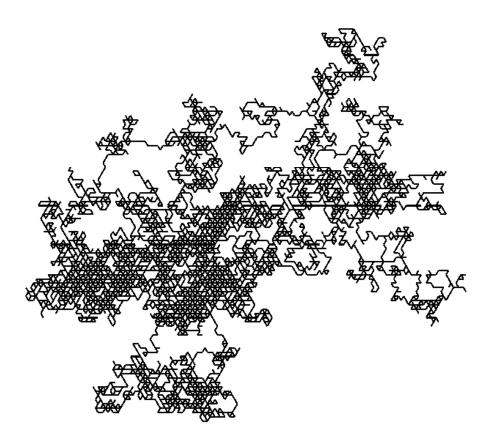
10 steps on triangular lattice



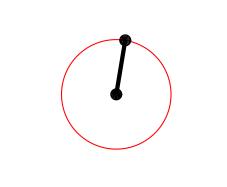
100 steps on triangular lattice

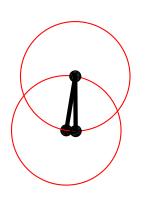


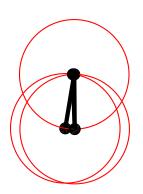
 $1000\ {\rm steps}$  on triangular lattice



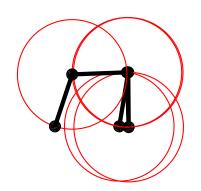
### 10000 steps on triangular lattice

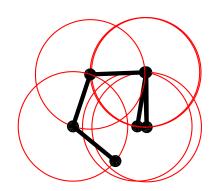


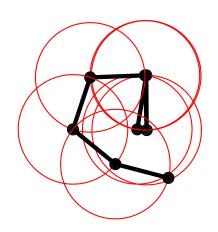


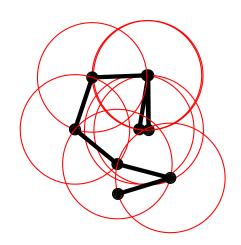


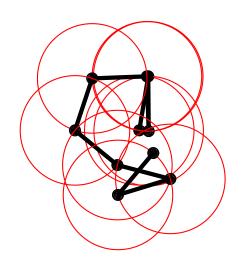


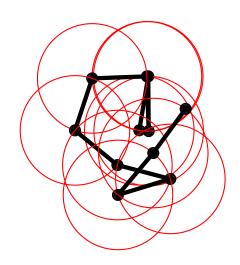


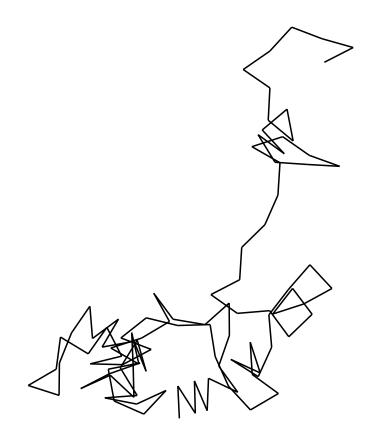




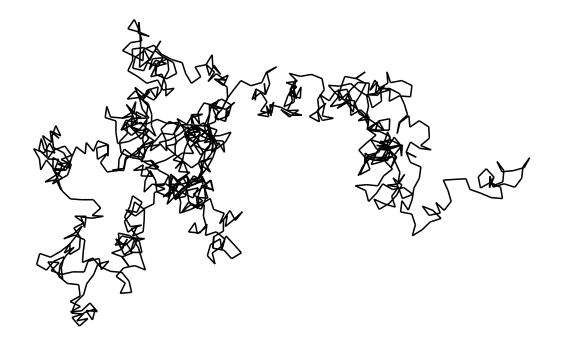




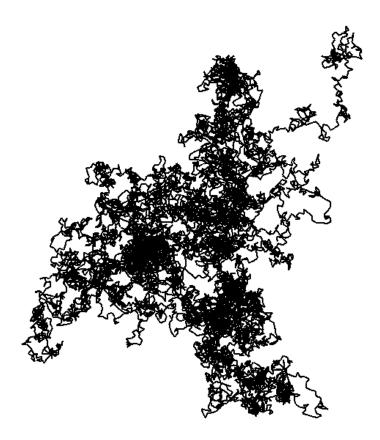




100 steps



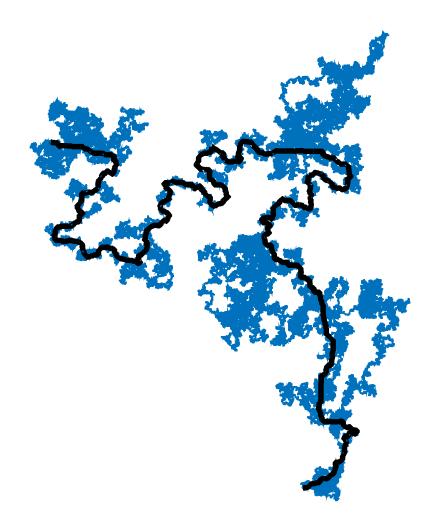
1000 steps



10000 steps

**Donsker's Invariance Principle:** Any iid random walk with mean zero and finite variance converges to Brownian motion.

#### PART II: SHORT PATHS IN THE BROWNIAN TRACE



Brownian motion = limit of rescaled random walks

It a "random continuous path" in plane.

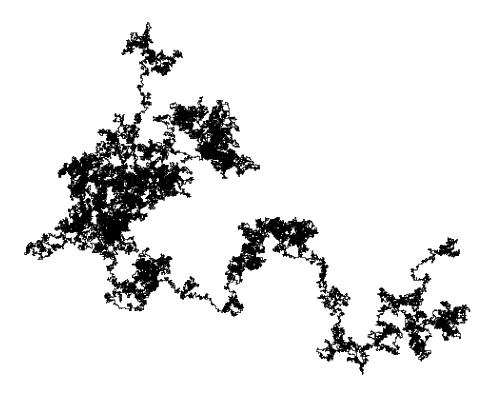
Run forever is dense in plane.

Run for unit time gives compact set = Brownian trace.

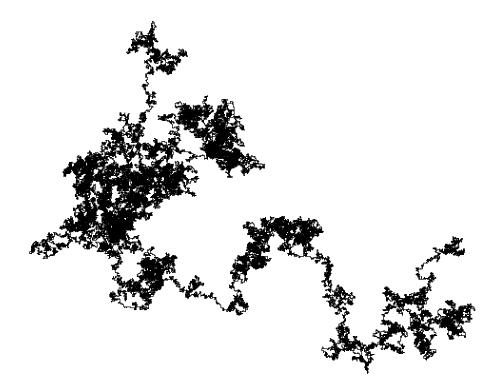
Zero area, dimension 2. Infinitely many complementary components.

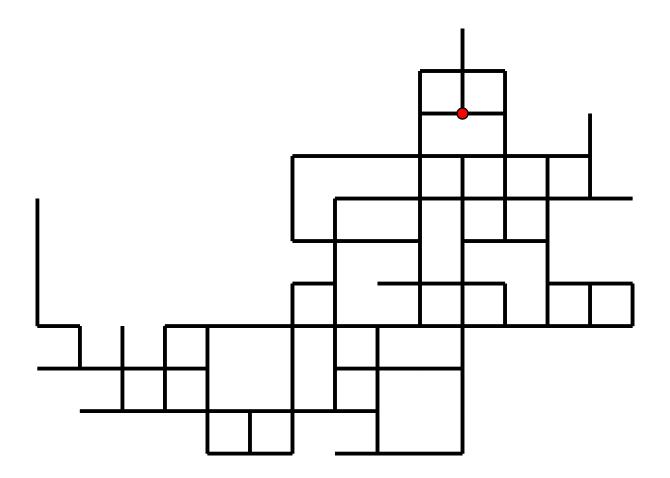


Robin Pemantle proved Brownian motion does not cover a line segment. **Question:** Does Brownian trace cover a rectifiable curve? Can points distance n apart be connected by path of length O(n)?

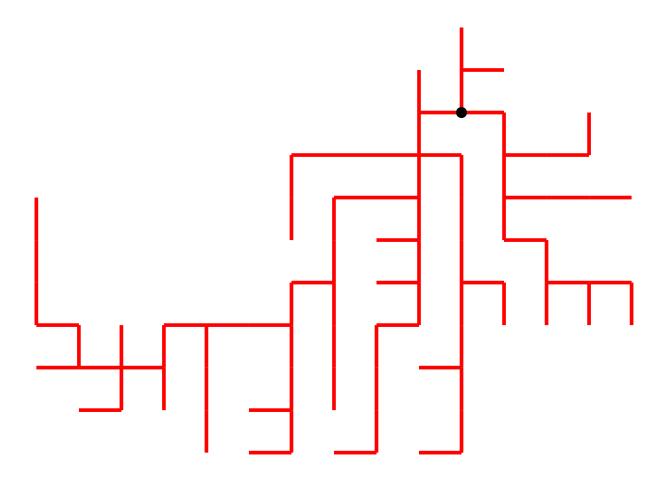


What about length  $O(n^d)$  for some d > 1 (dimension of path). Frontiers have dimension d = 4/3 = 1.333... (Lawler et. al.). Trace contains curves of dimension d = 5/4 = 1.25 (Dapeng Zhan).

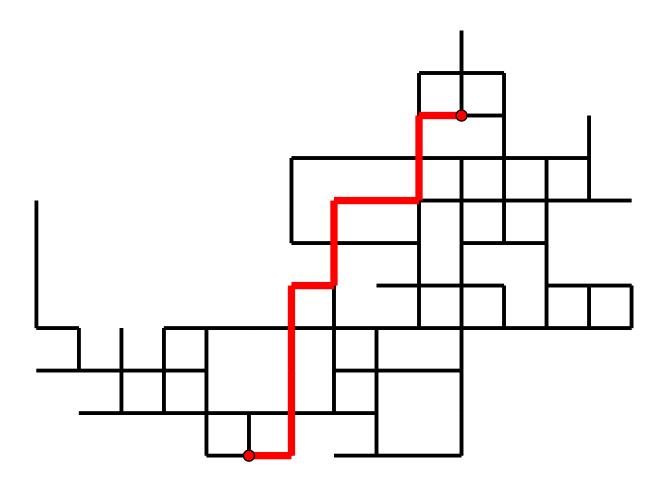




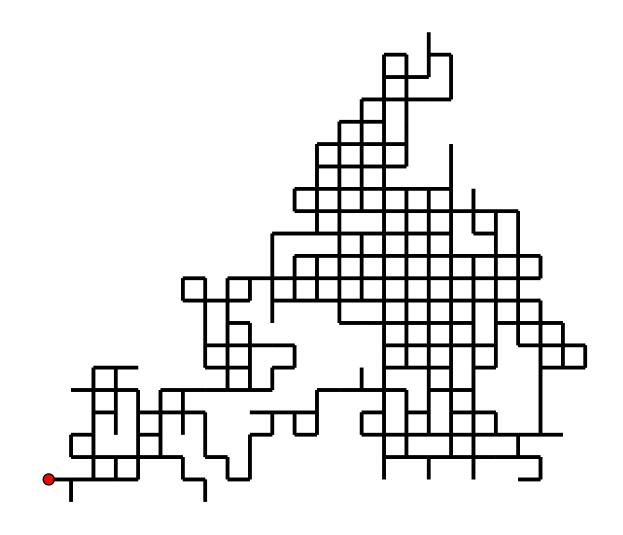
200 step random walk.



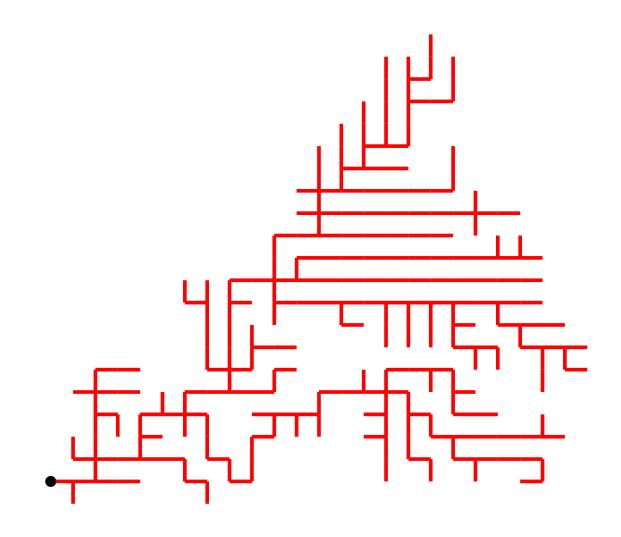
Minimal distance rooted spanning tree.



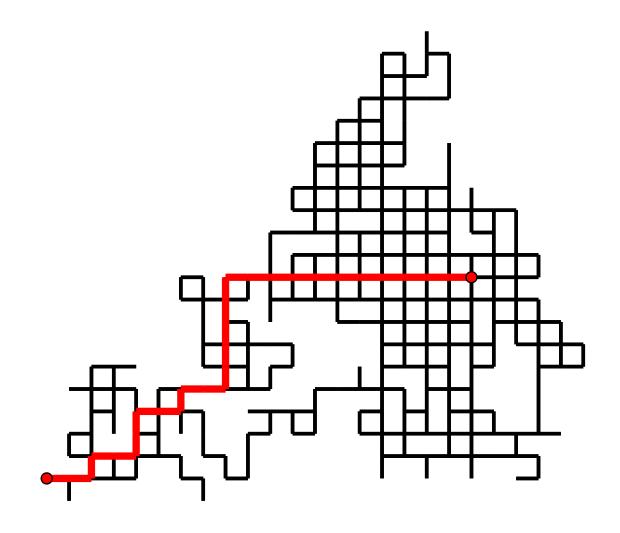
A shortest path from 0 to  $\sqrt{n}/2$ .



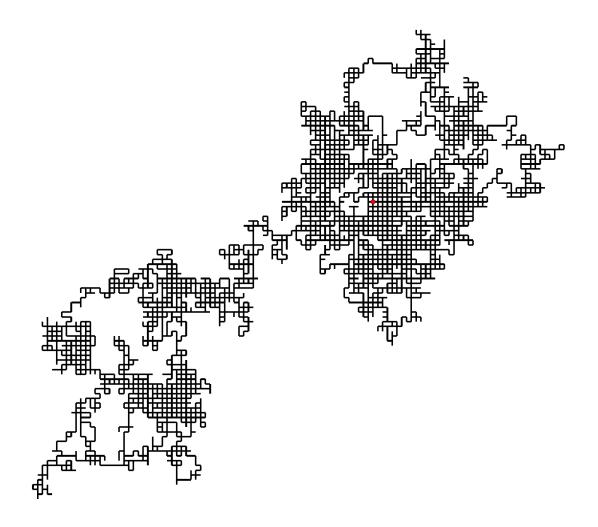
1000 step random walk.



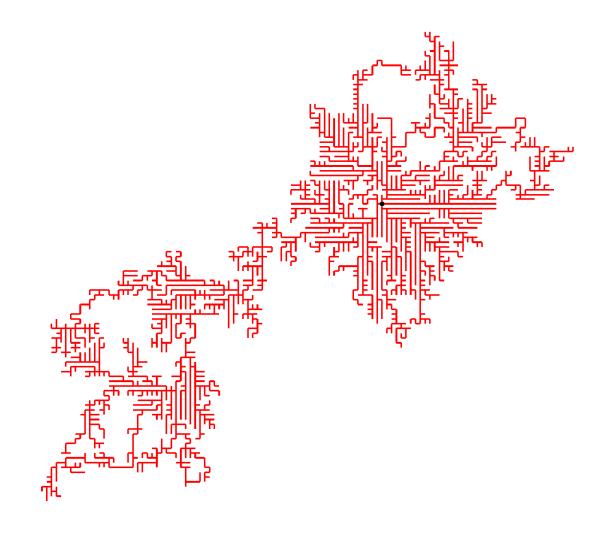
Minimal distance rooted spanning tree.



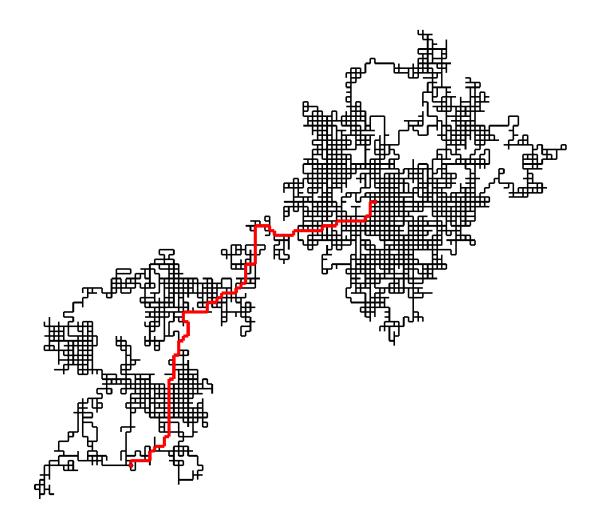
A shortest path from 0 to  $\sqrt{n}/2$ .



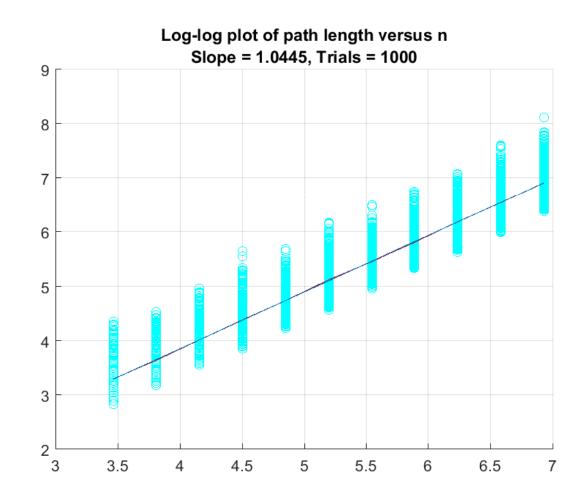
10000 step random walk.



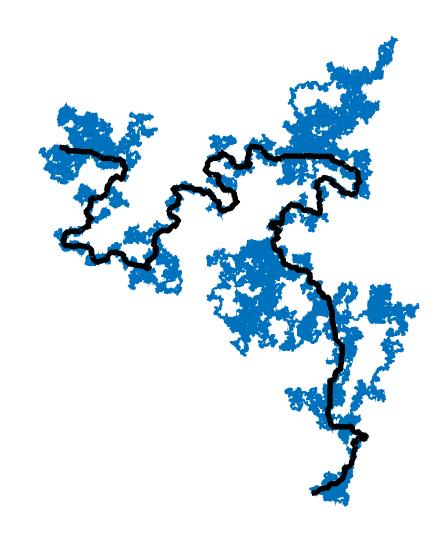
### Minimal distance spanning tree, wrt to origin.

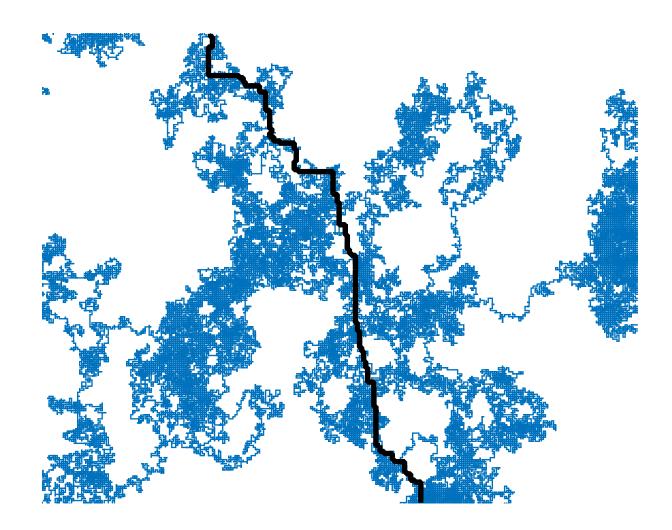


Does this curve converge to a fractal as  $n \to \infty$ ?.

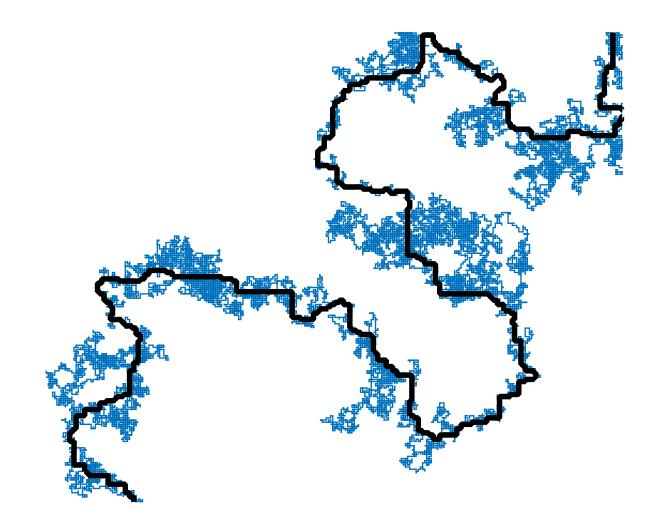


Log-log plot of graph distance 0 to  $\{|z| = \frac{1}{2}\sqrt{n}\}.$ 

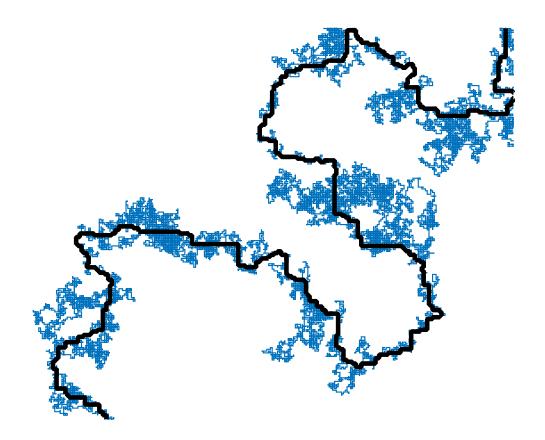




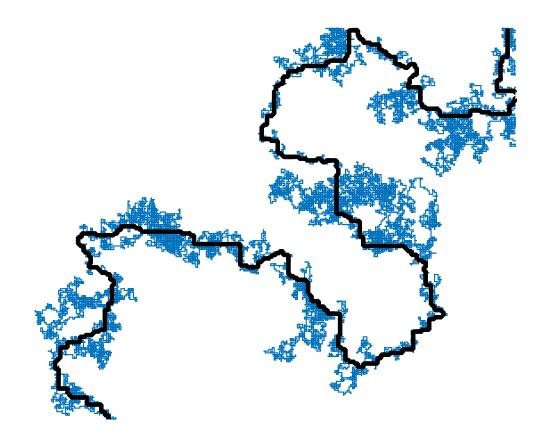
### Paths can be straight where trace is dense.



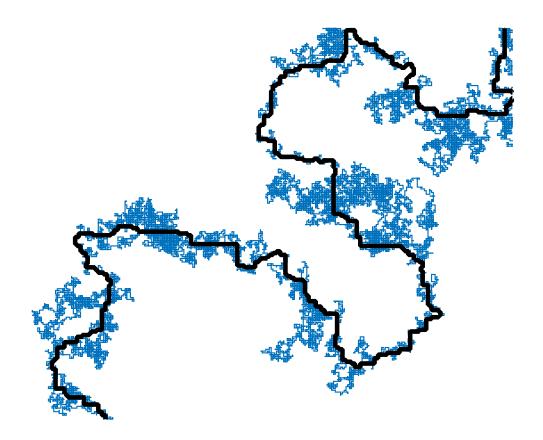
Paths wiggle more when trapped between components.



Can any two components be separated by a rectifiable curve?

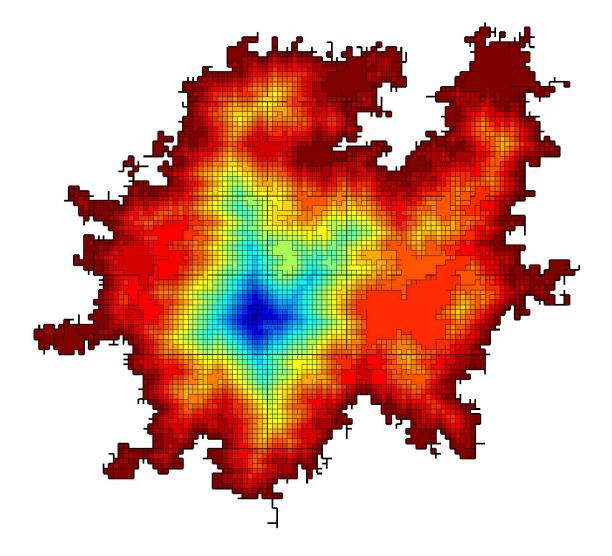


### Is the intersection of two boundaries rectifiable?



Intersection of component boundaries is empty or dimension 3/4. Self-similar sets of dimension < 1 are rectifiable.

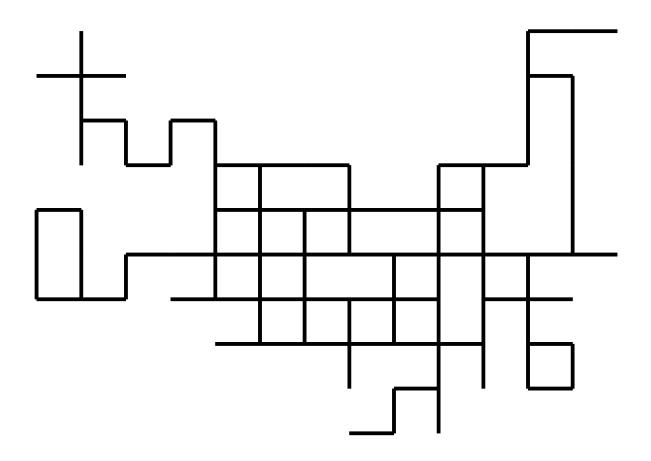
### PART III: THE GRAPH OF COMPLEMENTARY COMPONENTS



Consider the complementary components of the Brownian trace as vertices of a graph, with two being adjacent if their boundaries overlap.

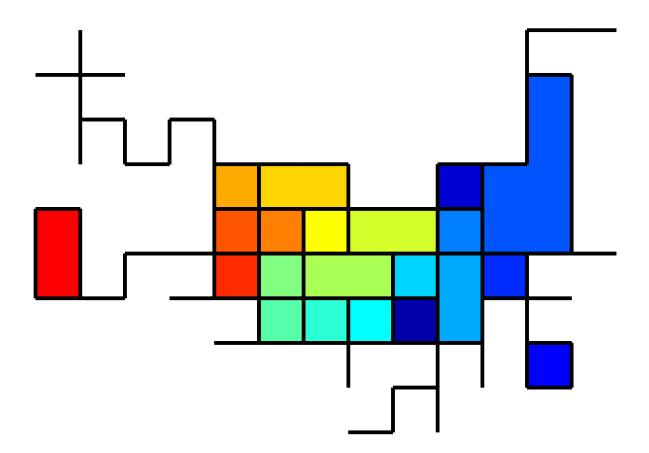
Wendelin Werner conjectured this graph is connected, i.e., any two components are connected by a path hitting the trace only finitely often.

What about a path that hits the trace countably often? Hits in a set of small Hausdorff dimension?



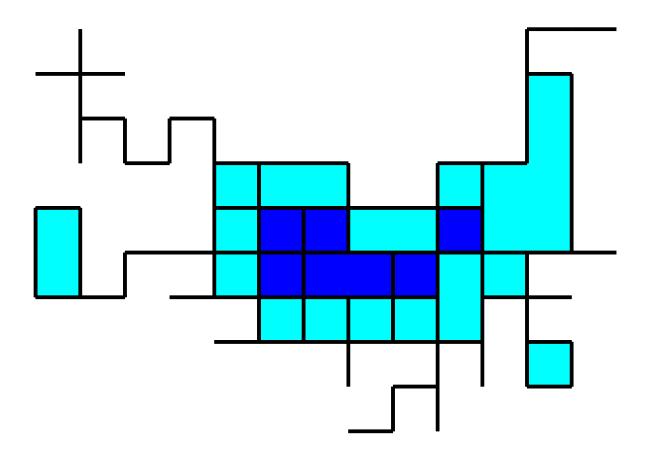
200 random steps on square grid.

N = 200, Number of components = 22



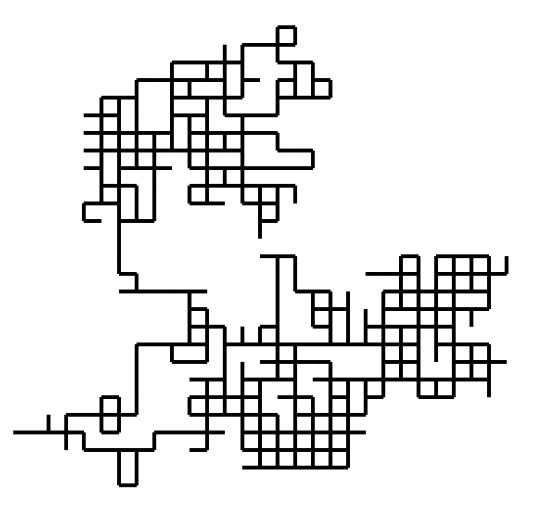
Components form a graph under edge adjacency.

N = 200, Number of components = 22, Depth = 2

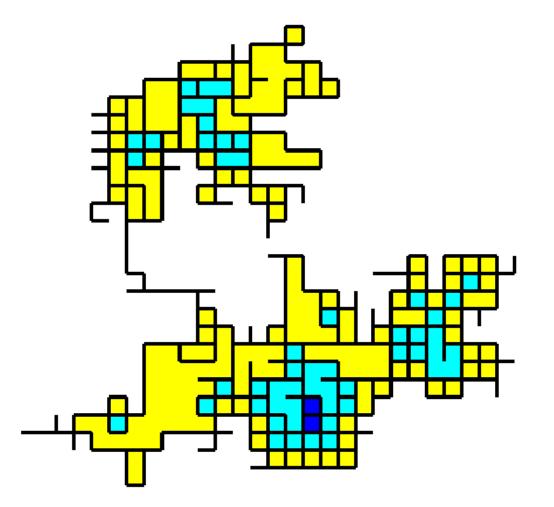


components colored by graph distance to outside.

N = 1000, The trace

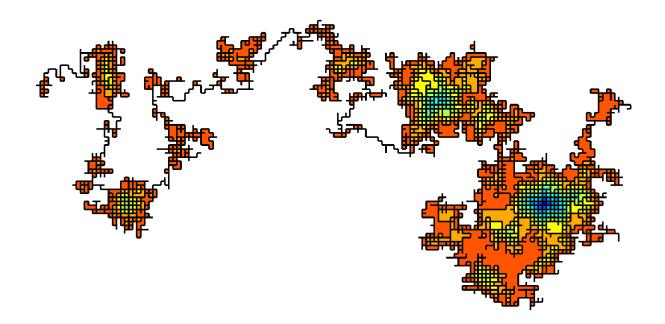


1000 random steps on square grid.

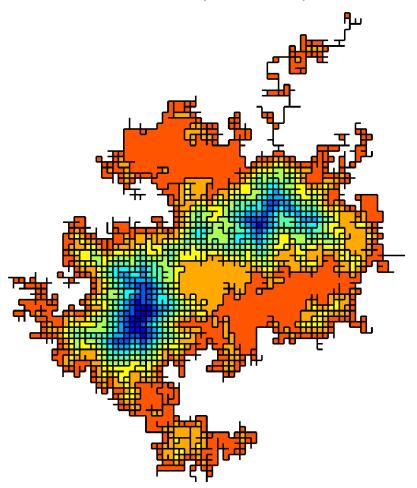


Components colored by graph distance to outer component.

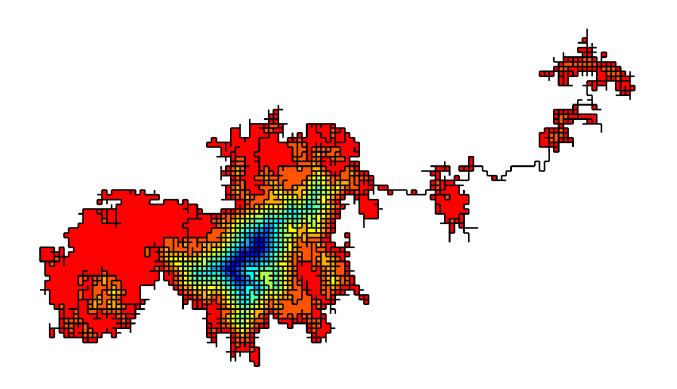
N = 10000, Number of components = 1132, Depth = 9

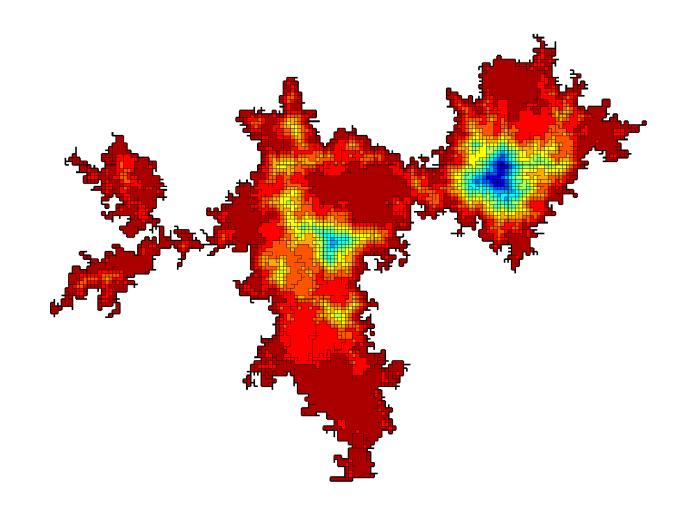


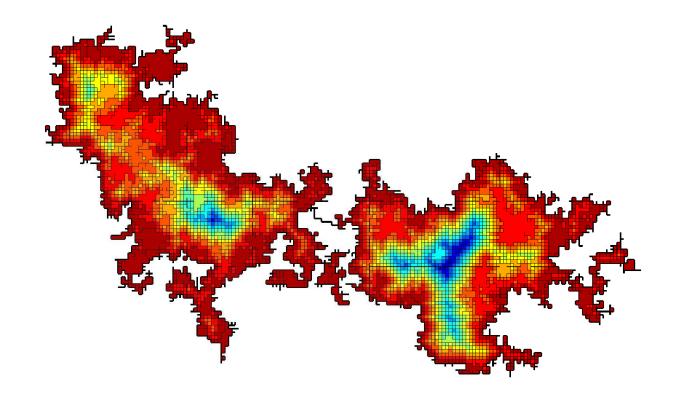
N = 10000, Number of components = 1147, Depth = 10



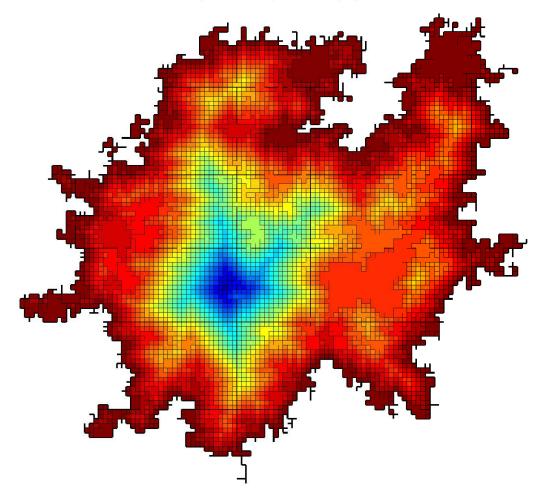
N = 10000, Number of components = 1034, Depth = 11



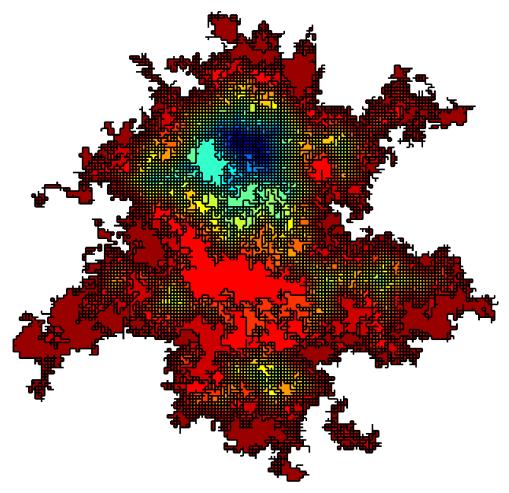




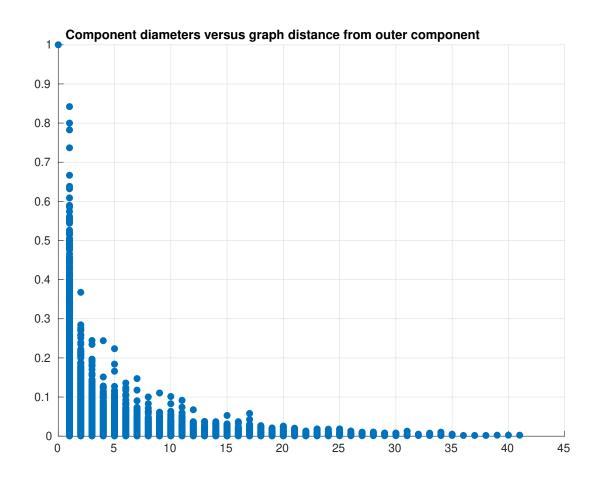
N = 40000, Number of components = 4019, Depth = 24



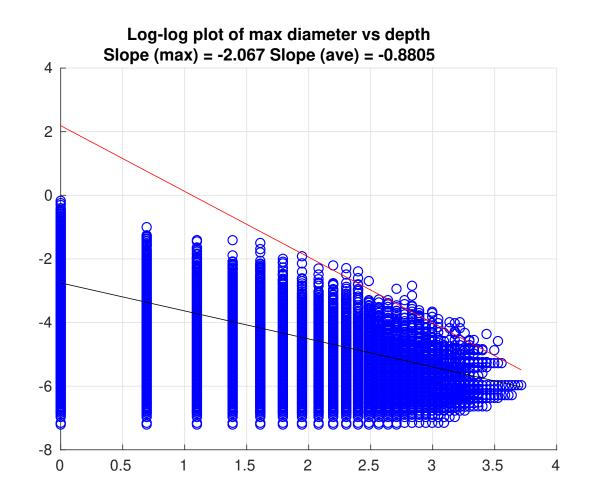
N = 50000, Number of components = 5548, Depth = 20



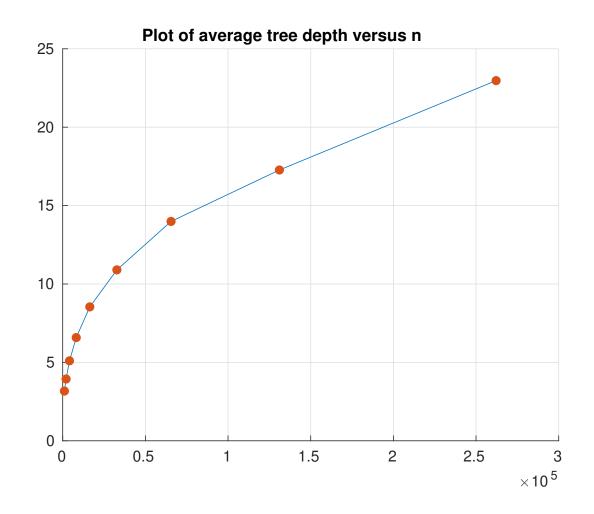
## 50,000 steps, 10%-component at depth 12 $\,$



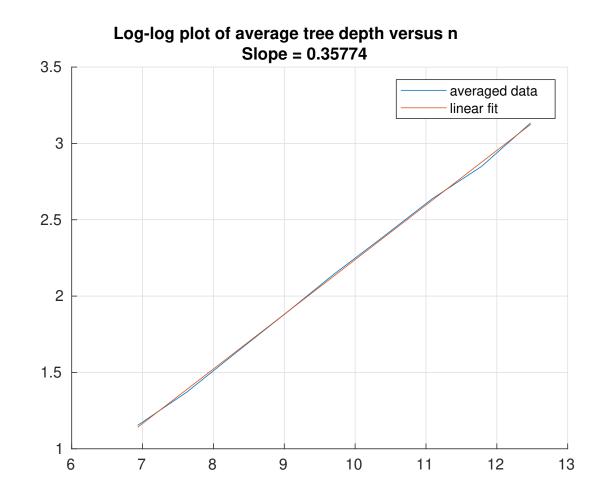
Component diameter versus graph distance from outer component.



Same plot in log-log coordinates.

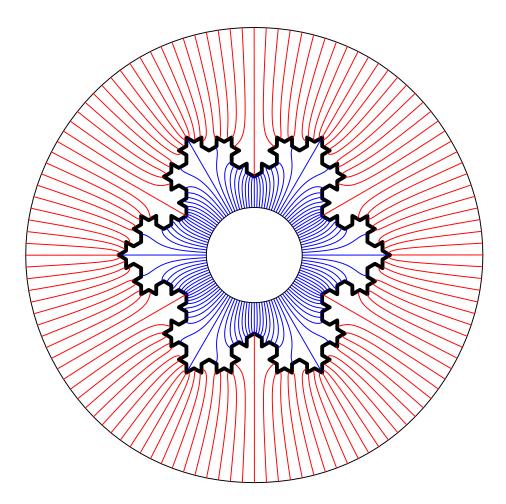


Growth of maximal graph distance (depth) to outer component.



Maximum distance from outer component looks like  $n^{.36}$ .

# PART IV: HARMONIC MEASURE AND CONFORMAL MAPS

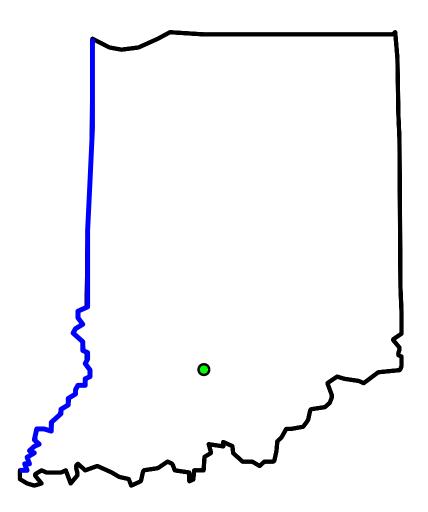




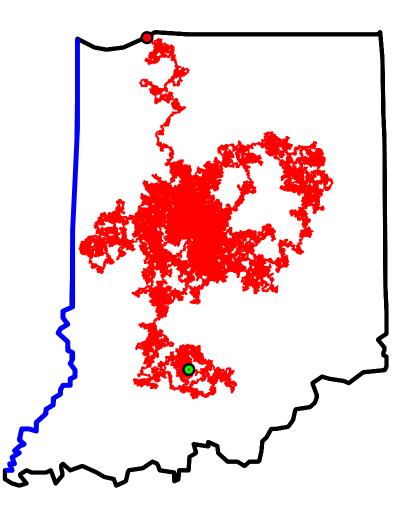
Suppose  $\Omega$  is a planar Jordan domain.



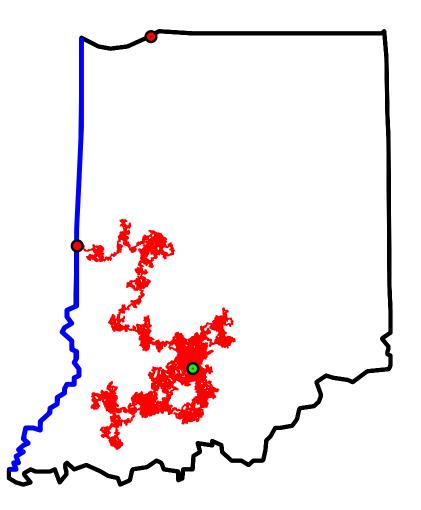
Let E be a subset of the boundary,  $\partial \Omega$ .



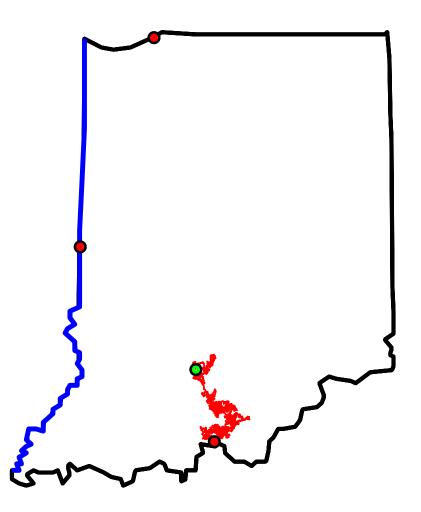
Choose an interior point  $z \in \Omega$ .



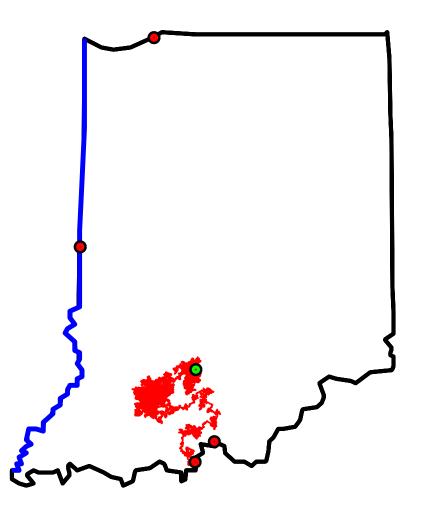
 $\omega(z, E, \Omega) =$ probability a particle started at z first hits  $\partial \Omega$  in E.



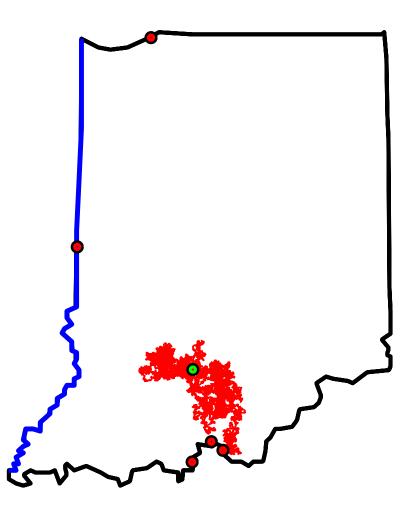
 $\omega(z, E, \Omega) =$ probability a particle started at z first hits  $\partial \Omega$  in E.



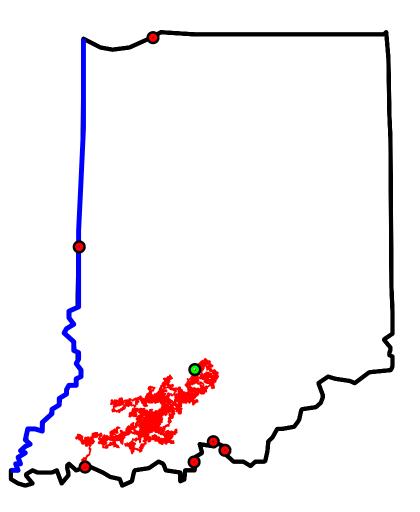
 $\omega(z, E, \Omega) =$ probability a particle started at z first hits  $\partial \Omega$  in E.



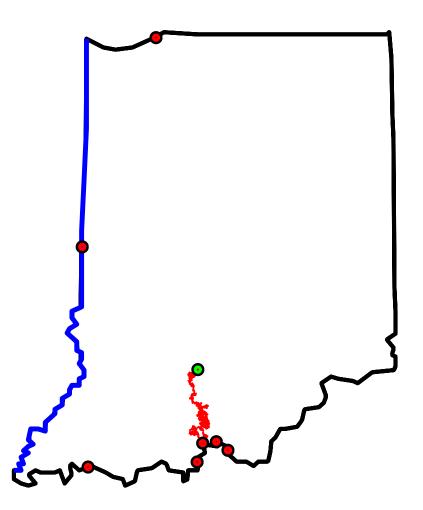
 $\omega(z, E, \Omega) =$ probability a particle started at z first hits  $\partial \Omega$  in E.



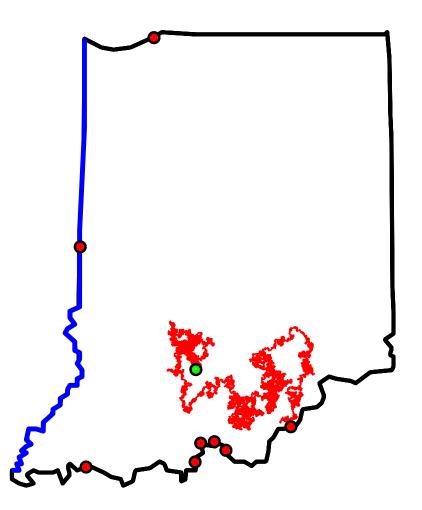
 $\omega(z, E, \Omega) =$ probability a particle started at z first hits  $\partial \Omega$  in E.



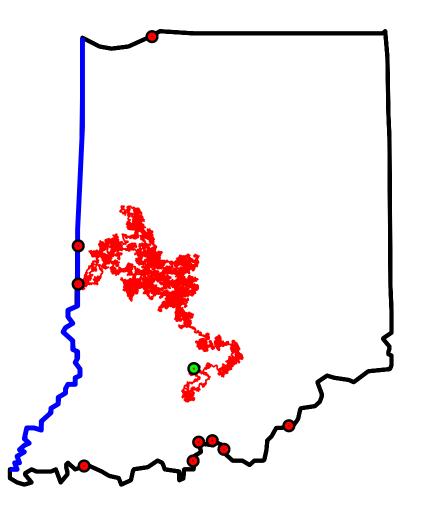
 $\omega(z, E, \Omega) =$ probability a particle started at z first hits  $\partial \Omega$  in E.



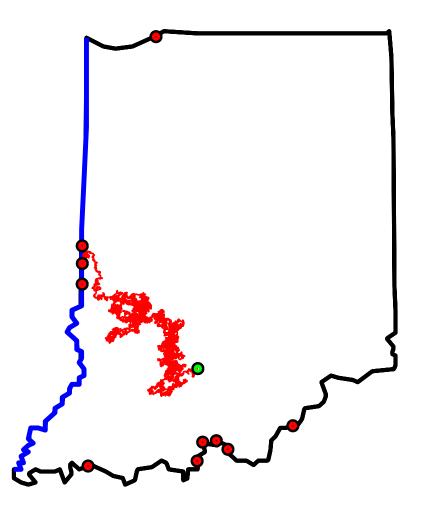
 $\omega(z, E, \Omega) =$ probability a particle started at z first hits  $\partial \Omega$  in E.



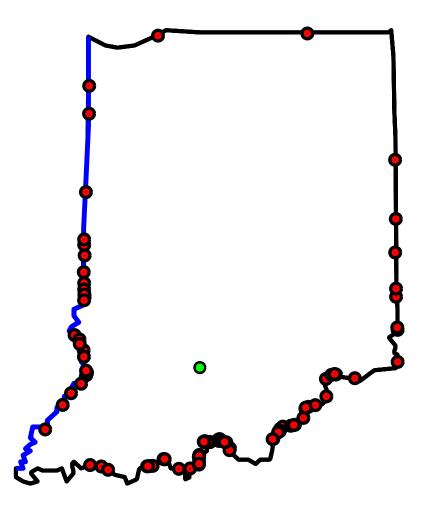
 $\omega(z, E, \Omega) =$ probability a particle started at z first hits  $\partial \Omega$  in E.



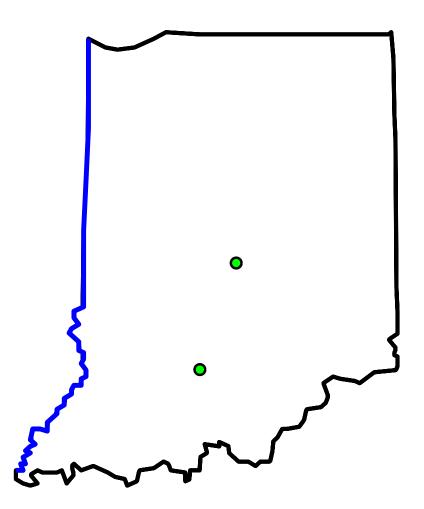
 $\omega(z, E, \Omega) =$ probability a particle started at z first hits  $\partial \Omega$  in E.



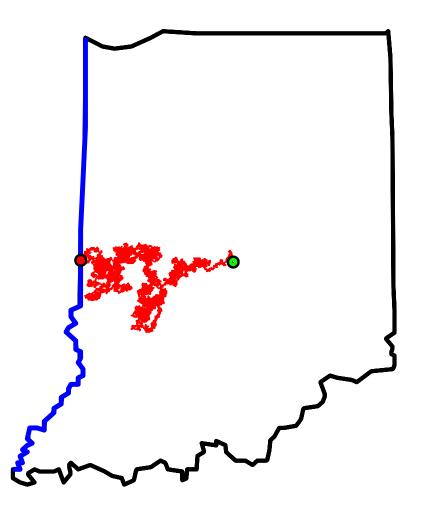
 $\omega(z, E, \Omega) =$ probability a particle started at z first hits  $\partial \Omega$  in E.

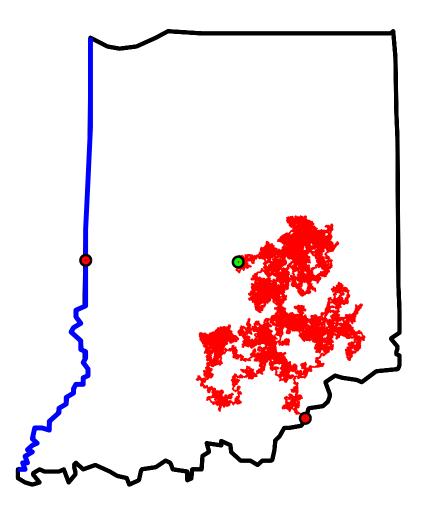


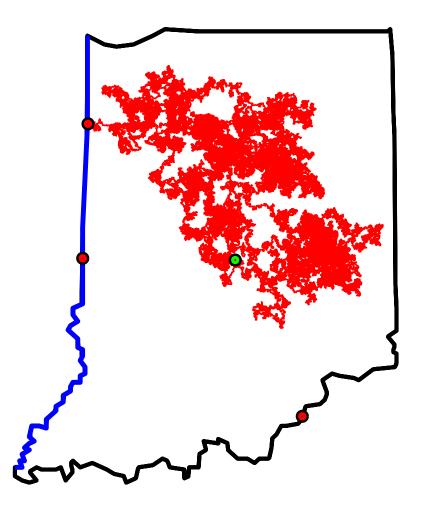
 $\omega(z_1, E, \Omega) \approx 30/100.$ 

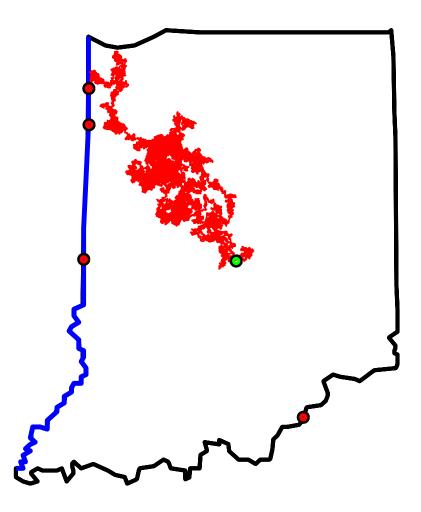


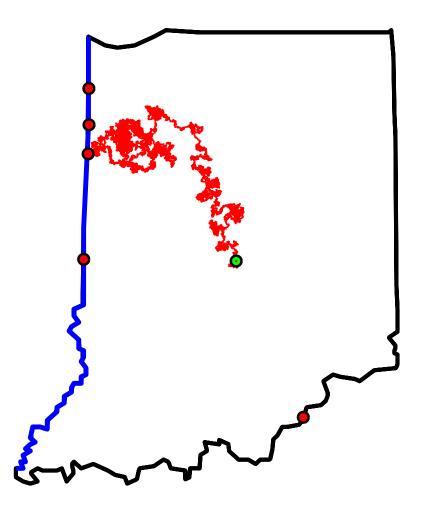
What if we move the starting point?

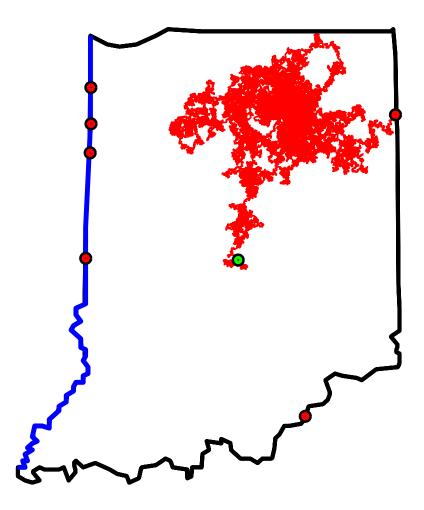


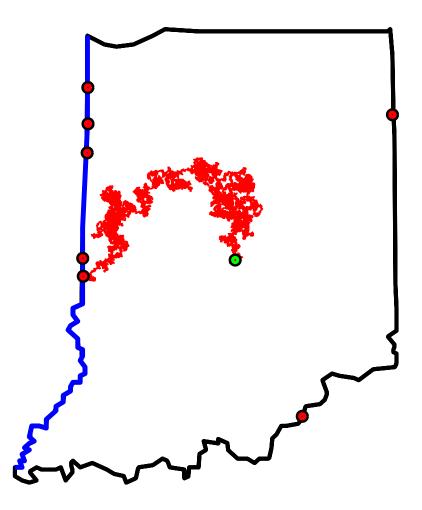


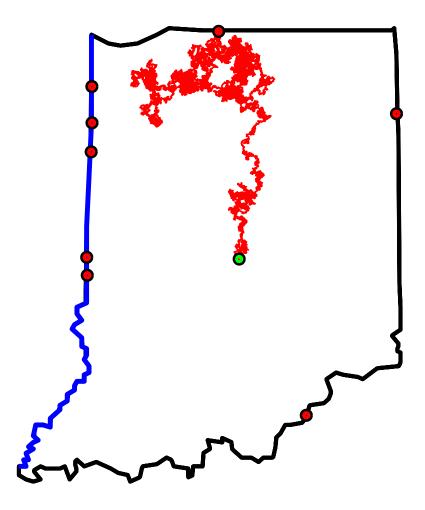


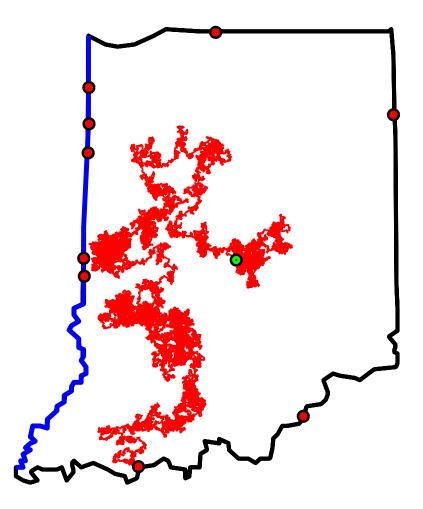


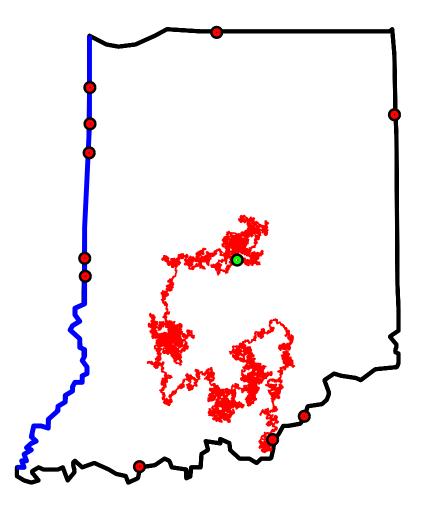


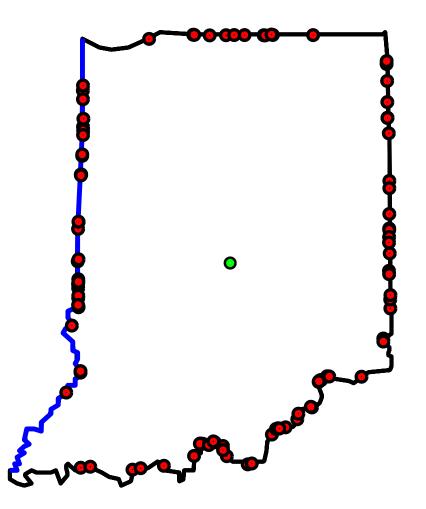




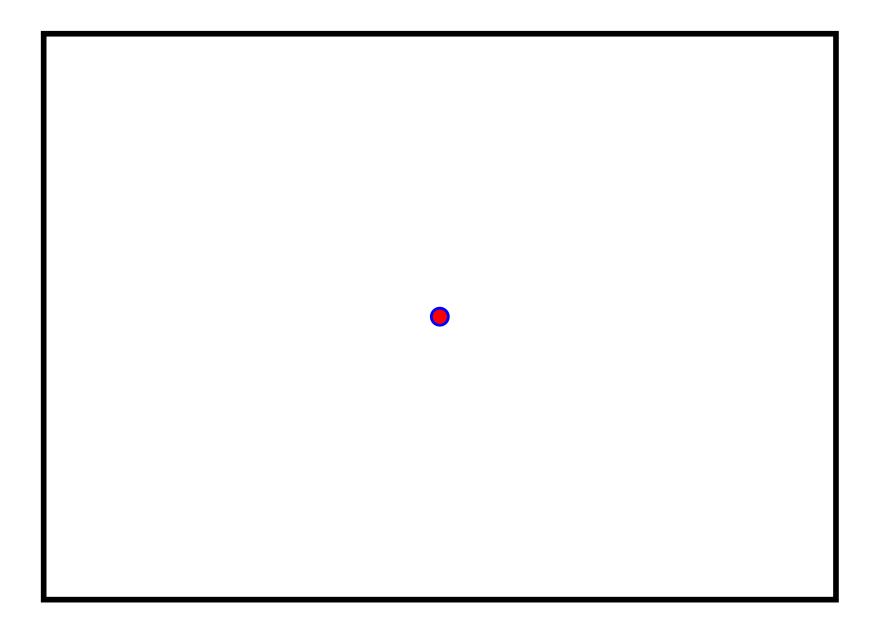


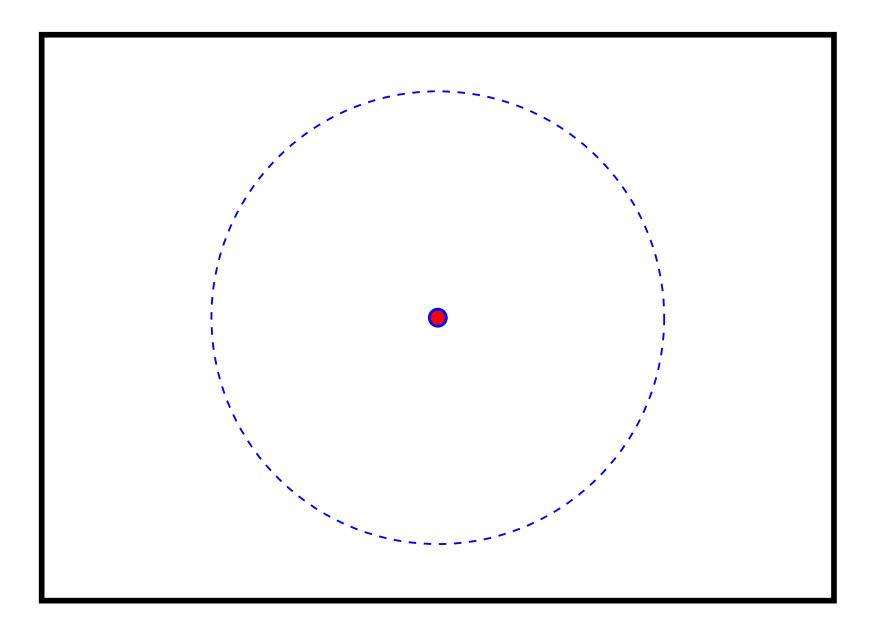


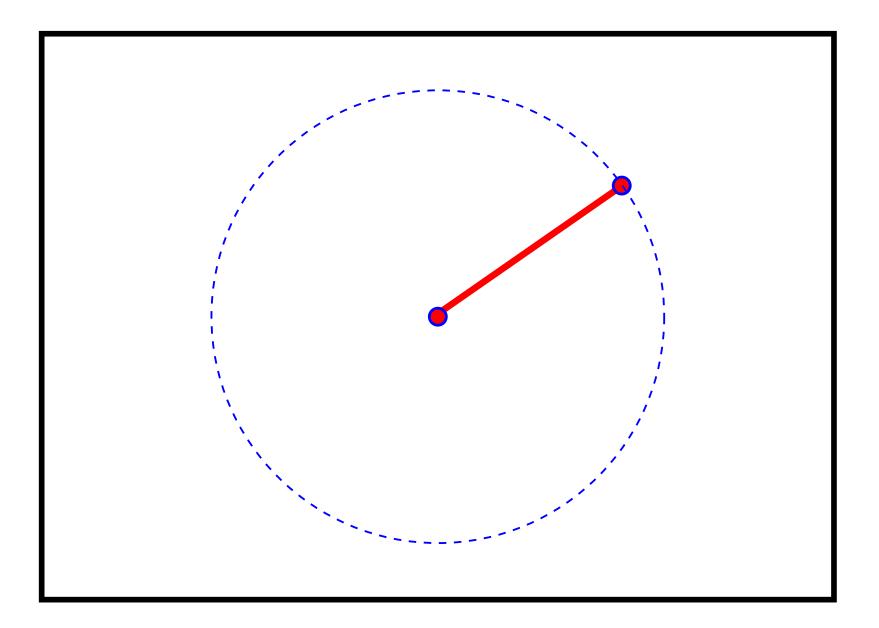


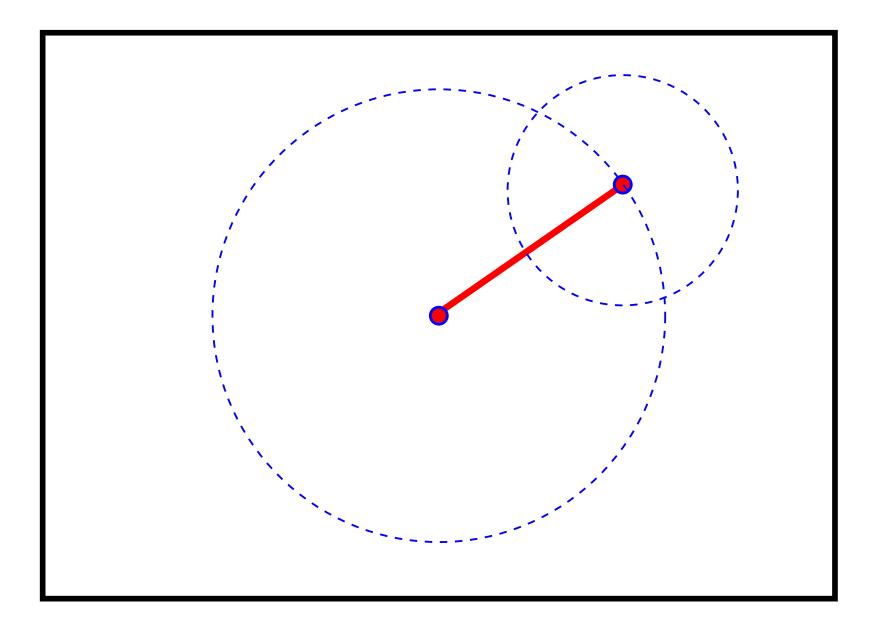


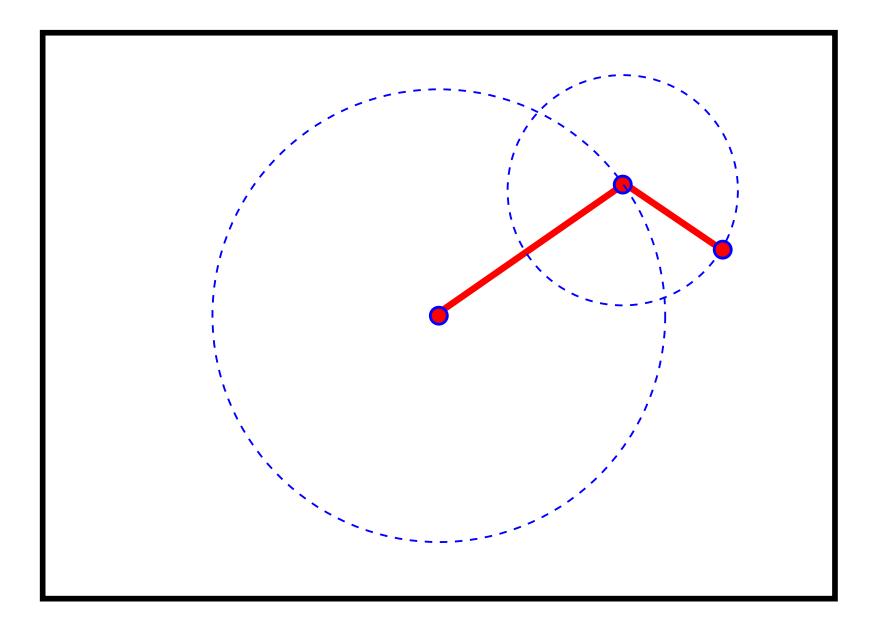
 $\omega(z_2, E, \Omega) \approx 34/100.$ 

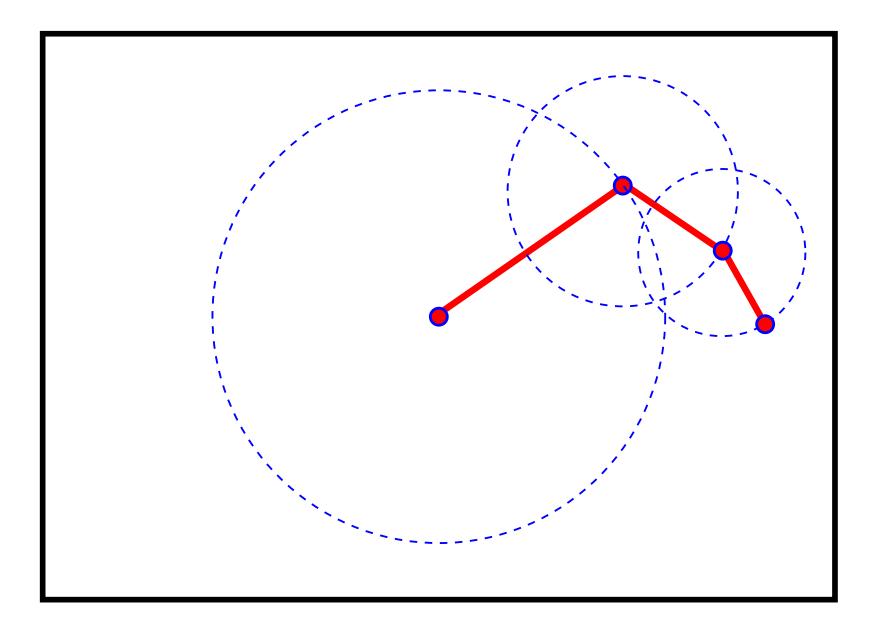


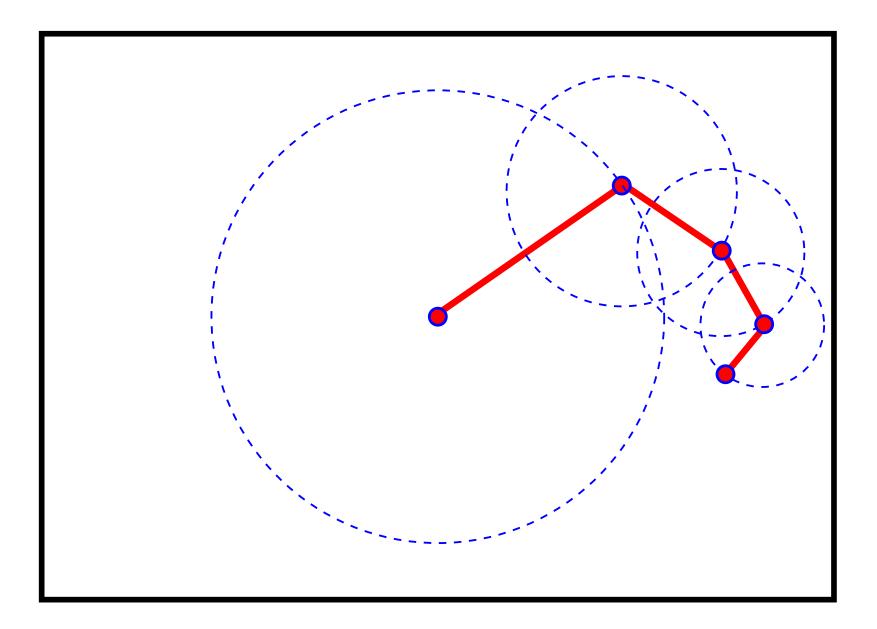


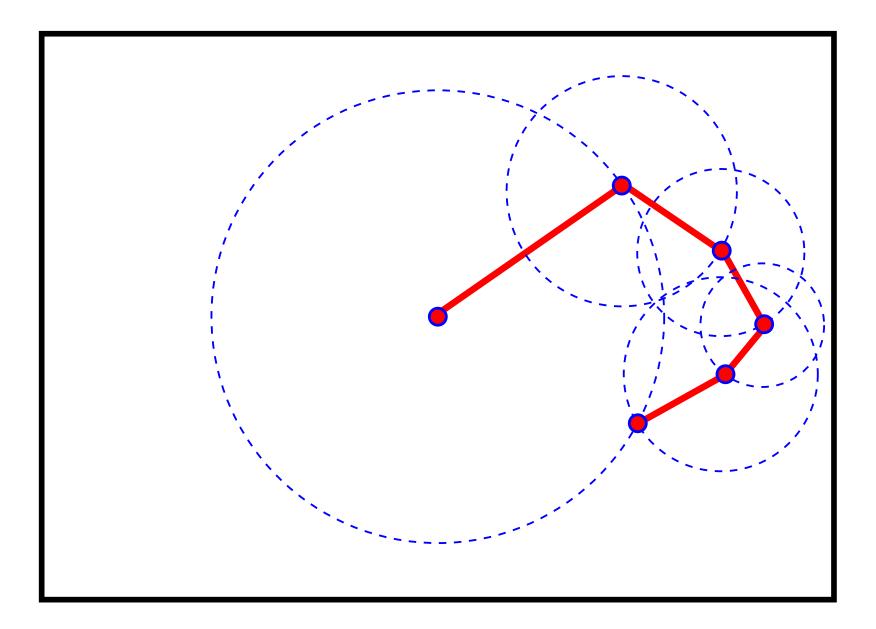


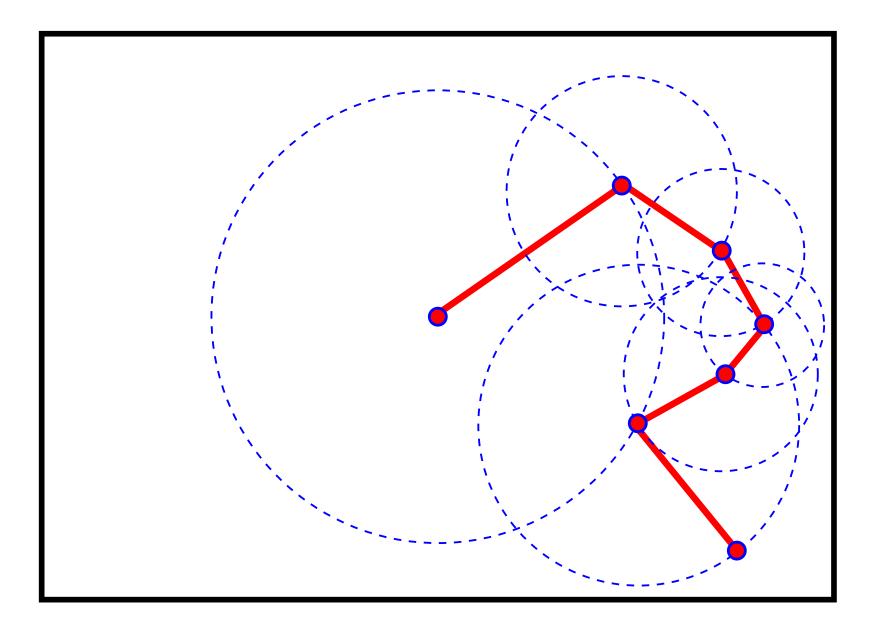


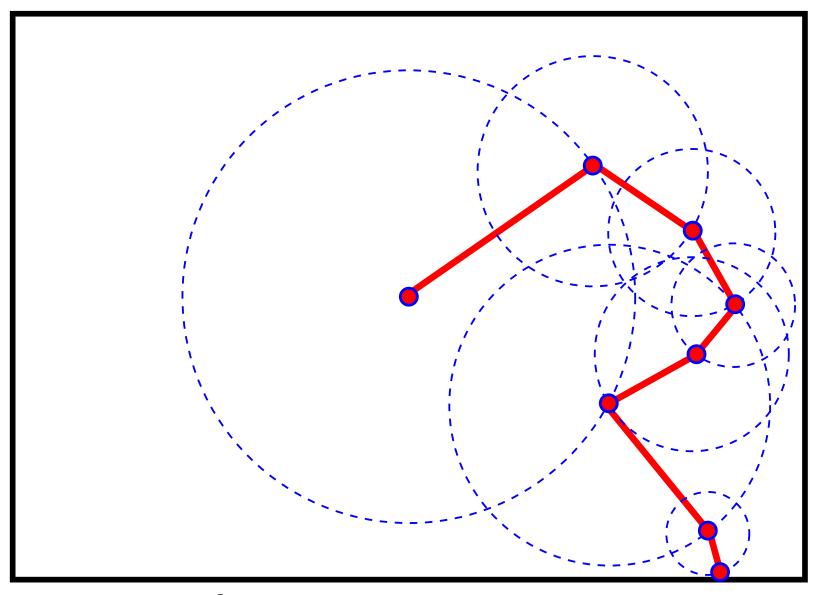




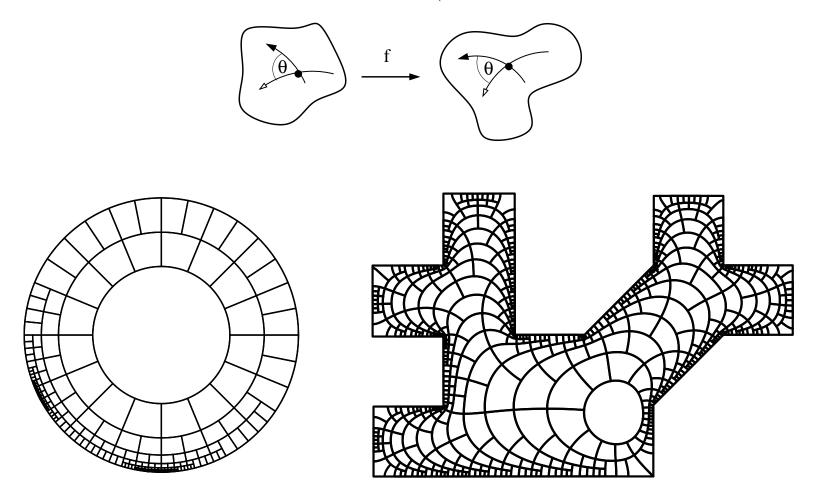








 $n^2$  versus log n to get within 1/n

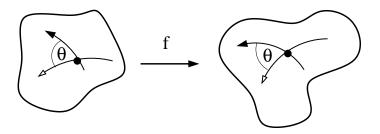


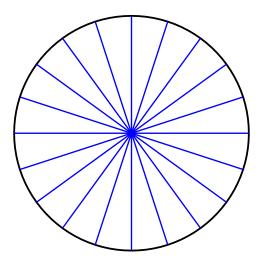


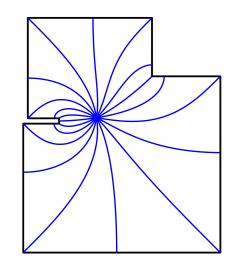
### Georg Friedrich Bernhard Riemann Stated RMT in 1851 thesis *a gloriously fertile originality* - Gauss

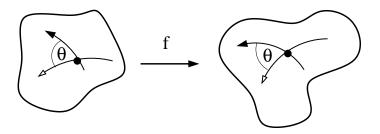


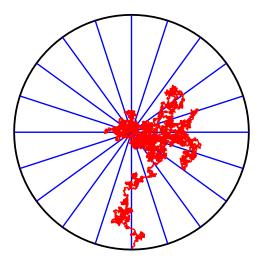
# William Fogg Osgood First proof of RMT, 1900

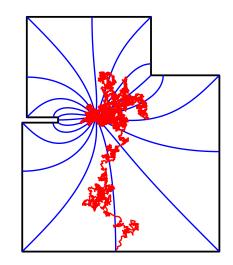


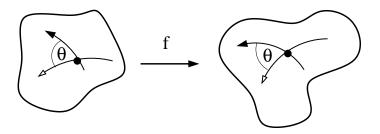


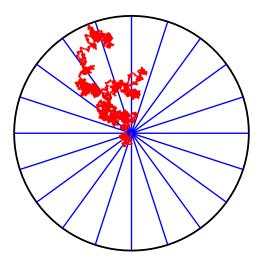


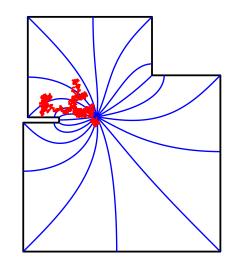


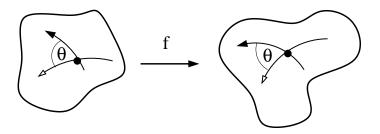


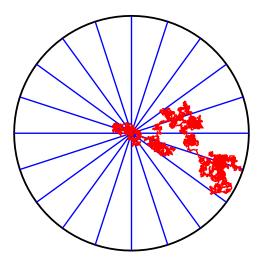


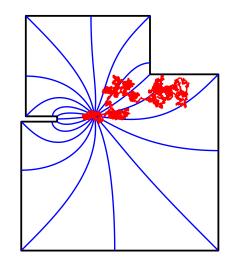


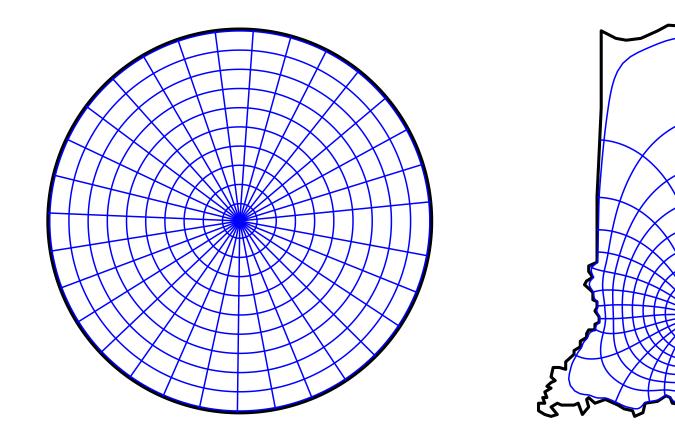








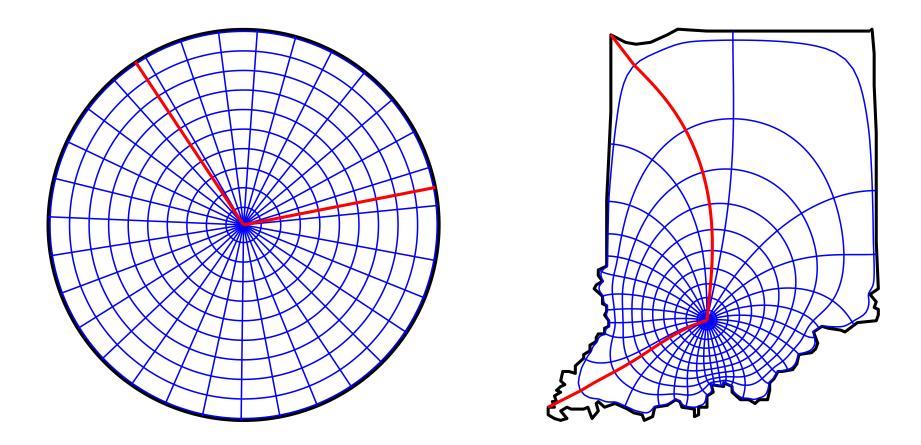




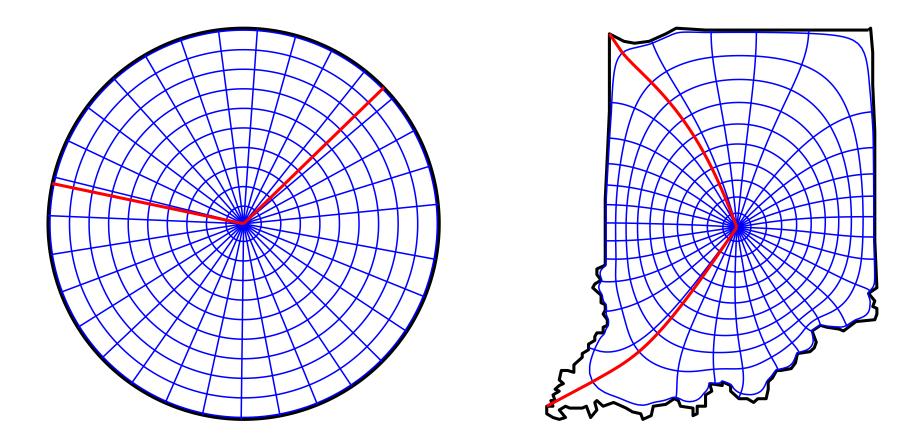
-

XIII

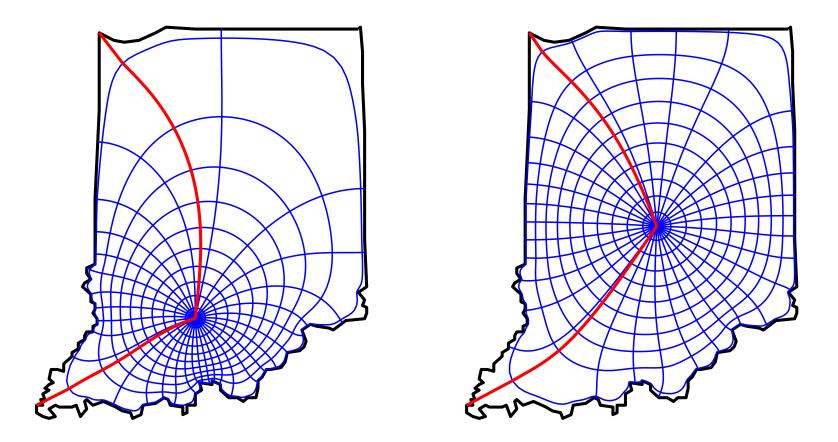
. .



harmonic measure = 0.31246



#### harmonic measure = 0.34431



### harmonic measure: 0.31246 versus 0.34431

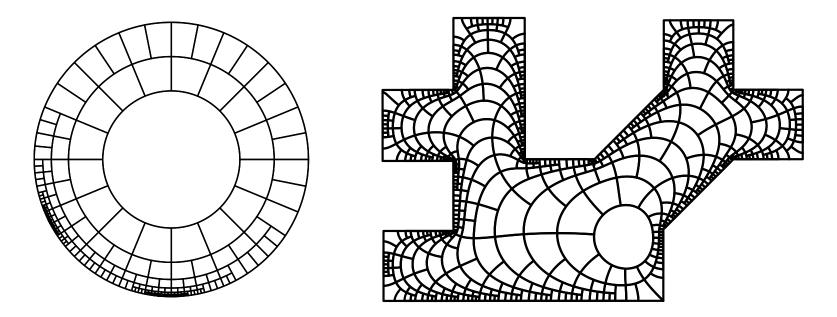
#### Some basic properties of harmonic measure:

 $\omega(z, E, \Omega)$  is the harmonic measure of  $E \subset \partial \Omega$ .

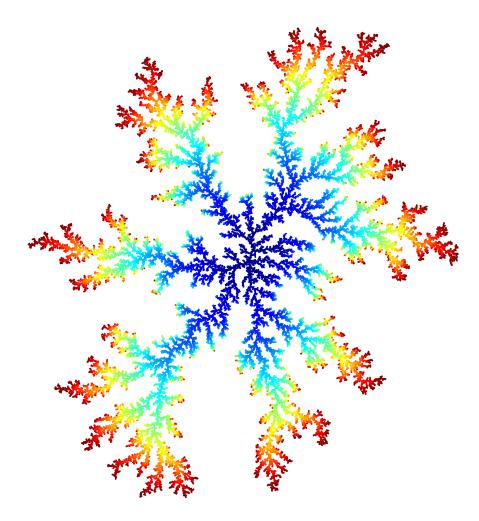
 $\omega$  is harmonic in z and  $0 \leq \omega \leq 1$ 

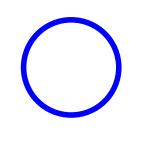
A harmonic function attaining its min or max is constant.

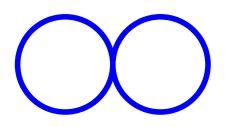
So a set E has positive measure for one point iff for all points.

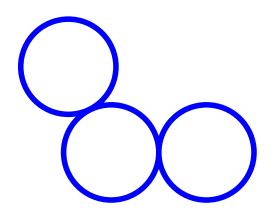


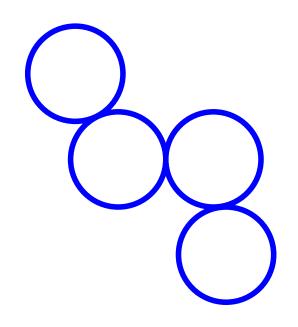
# PART V: DIFFUSION LIMITED AGGREGATION (DLA)

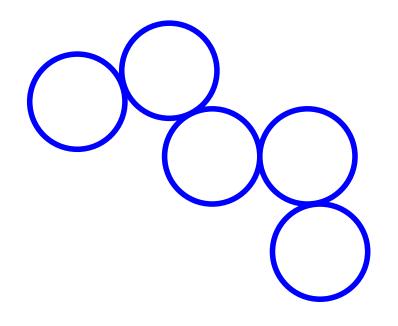


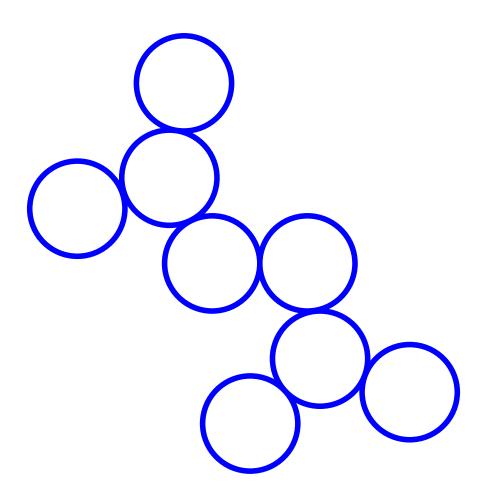


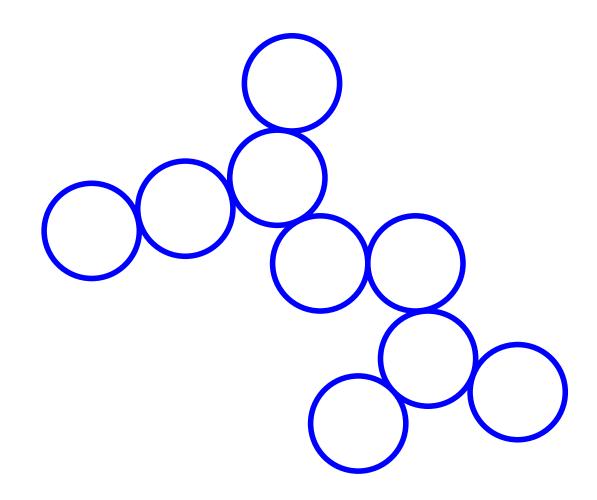


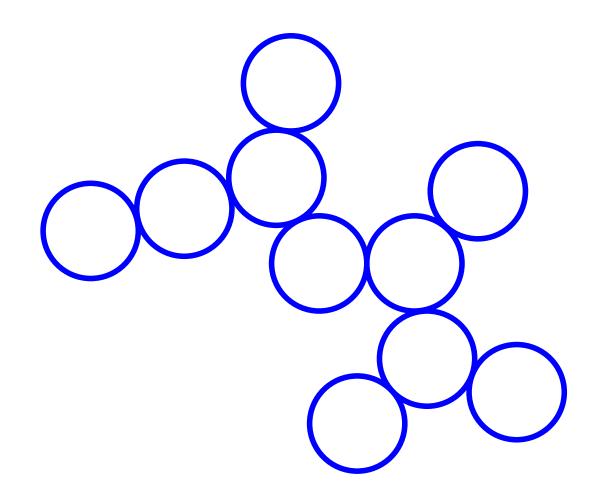


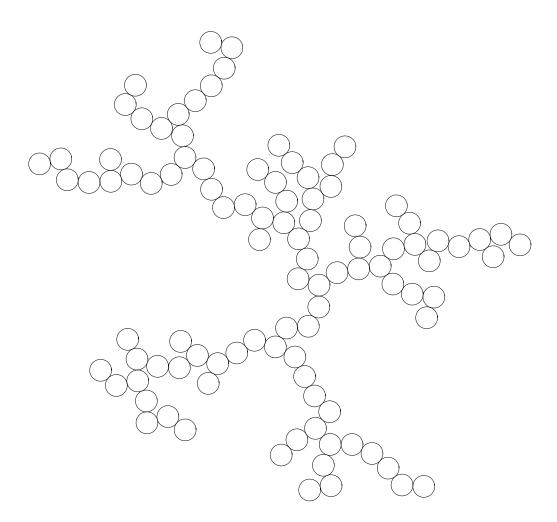


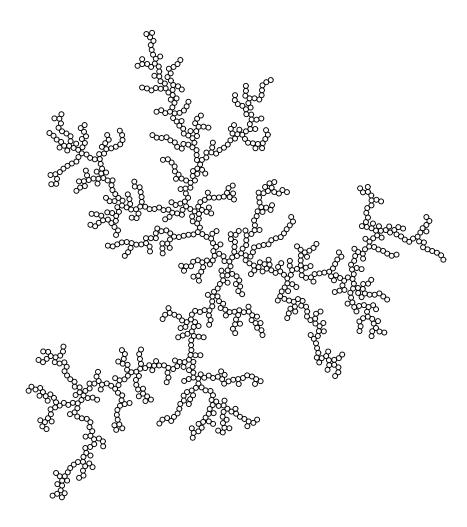


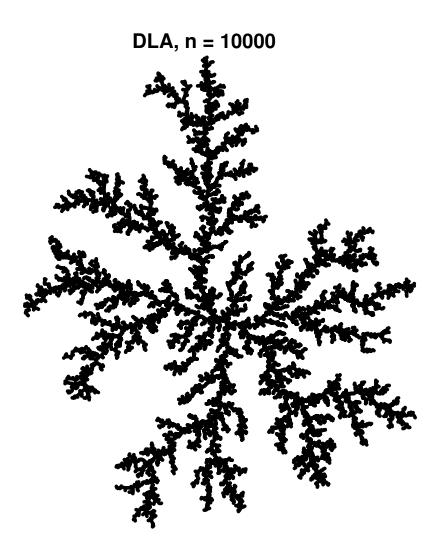


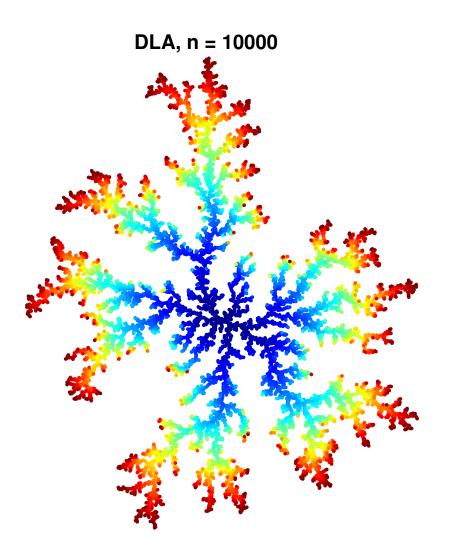


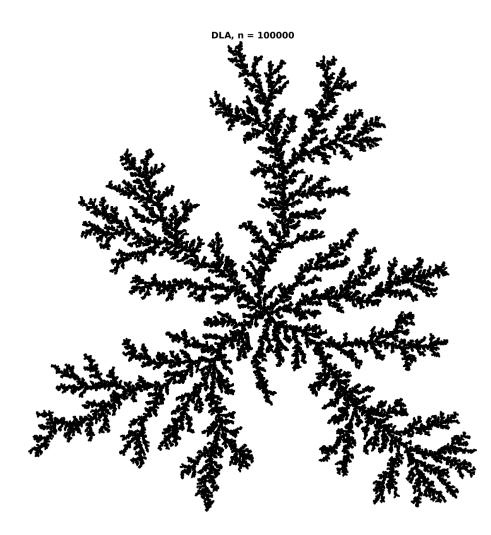


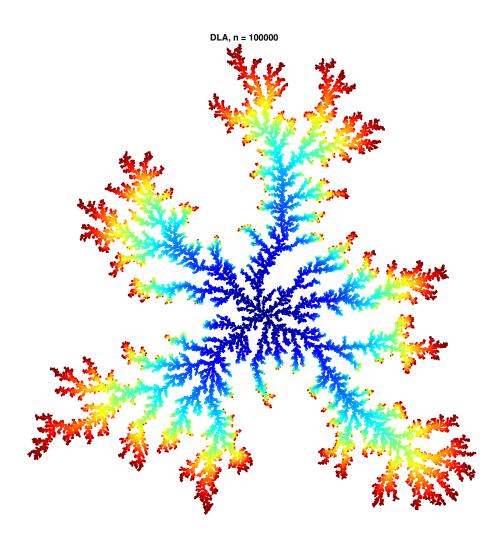




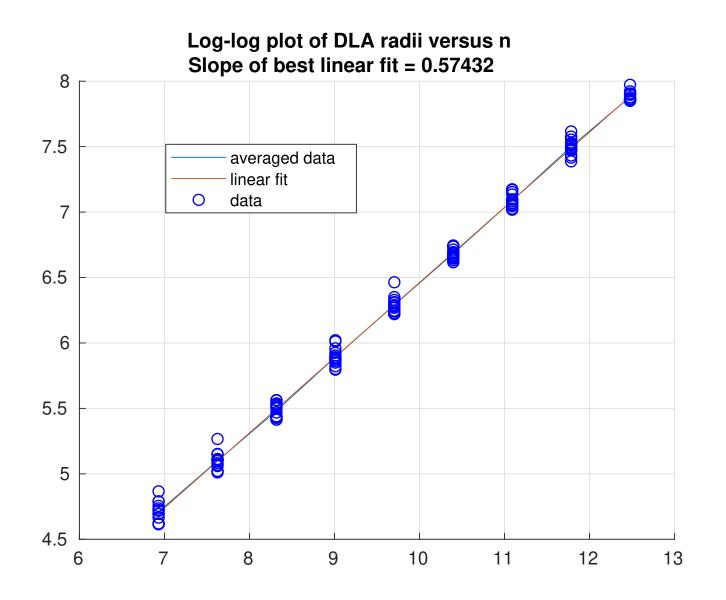




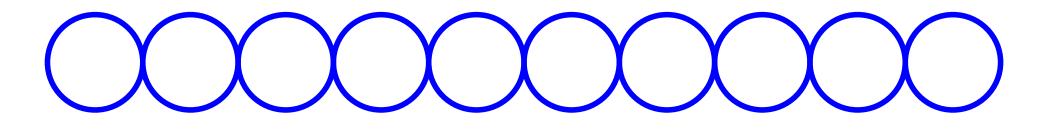




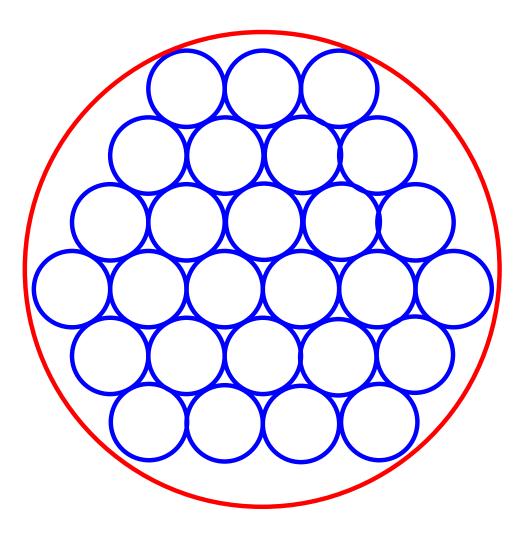
How fast does the diameter grow?



Numerical experiment for growth rate.



Trivial upper bound is O(n).



Trivial lower bound is  $\Omega(\sqrt{n})$ .

Theorem (Kesten): diam $(DLA(n)) = O(n^{2/3})$ .

Theorem (Kesten): diam(DLA(n)) =  $O(n^{2/3})$ .

Equivalent: DLA takes  $\gtrsim m^{3/2}$  steps to exit ball of radius m.

Suppose current radius is m. How long to reach 2m?

**Theorem (Beurling):** If K is connected and has diameter R, the harmonic measure of  $D(x, 1) \cap K$  is  $\leq C/\sqrt{R}$ .

harmonic measure = first hitting distribution of Brownian motion

Proved using conformal maps, extremal length.

**Theorem (Beurling):** If K is connected and has diameter R, the harmonic measure of  $D(x, 1) \cap K$  is  $\leq C/\sqrt{R}$ .

 $\Rightarrow$  Given a unit disk in a cluster of diameter m, it takes about  $\sqrt{m}$  attempts to attach a new disk it.

 $\Rightarrow$  It takes  $m \cdot \sqrt{m}$  attempts to grow a path of length m.

**Theorem (Beurling):** If K is connected and has diameter R, the harmonic measure of  $D(x, 1) \cap K$  is  $\leq C/\sqrt{R}$ .

 $\Rightarrow$  Given a unit disk in a cluster of diameter m, it takes about  $\sqrt{m}$  attempts to attach a new disk it.

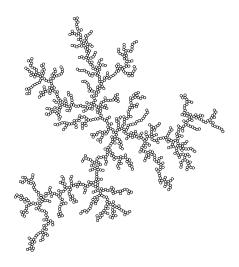
 $\Rightarrow$  It takes  $m \cdot \sqrt{m}$  attempts to grow a path of length m.

 $\Rightarrow$  It takes time  $m^{3/2}$  to grow from m to 2m.

Beurling's theorem is sharp when K is line segment. Since DLA never looks like a line segment,

- $\Rightarrow$  smaller estimate for harmonic measure,
- $\Rightarrow$  longer waiting time to cross disk,
- $\Rightarrow$  better upper bound for the diameter of DLA.

How to make this precise? Prove DLA is not a "line segment".



Amazingly, there is no known better lower bound than the trivial  $\sqrt{n}$ .

Conjecture: Almost surely,

$$\lim_{n \to \infty} \frac{\operatorname{diam}(\mathrm{DLA}(n))}{\sqrt{n}} = \infty.$$

Amazingly, there is no known better lower bound than the trivial  $\sqrt{n}$ .

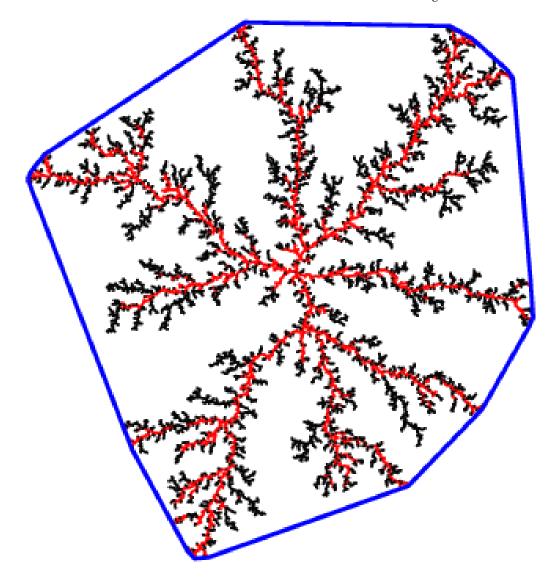
Conjecture: Almost surely,

$$\lim_{n \to \infty} \frac{\operatorname{diam}(\mathrm{DLA}(n))}{\sqrt{n}} = \infty.$$

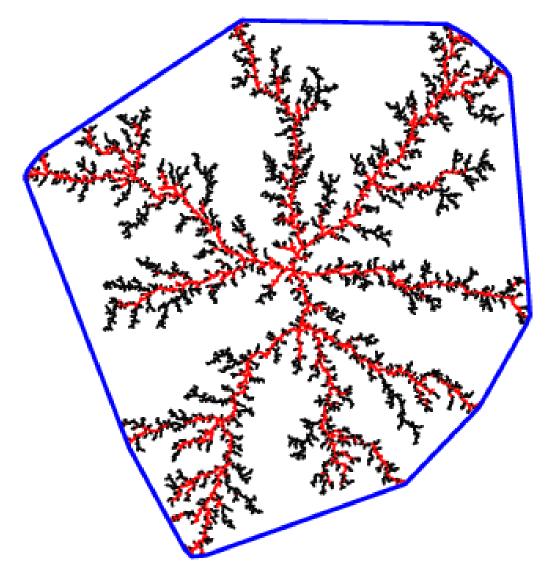
If DLA(n) is roughly a disk of radius  $\sqrt{n}$  then any boundary disk is hit with probability  $\simeq 1/\sqrt{n}$ , which gives which gives the trivial lower bound.

For non-trivial lower bound, we need to show there are points that get hit with probability  $\gg n^{-1/2}$ .

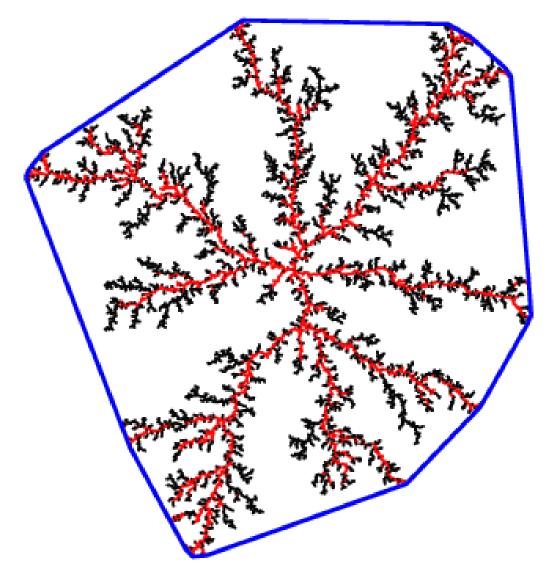
Consider convex hull of the DLA cluster. What is the harmonic measure of the disks that touch the convex hull boundary?

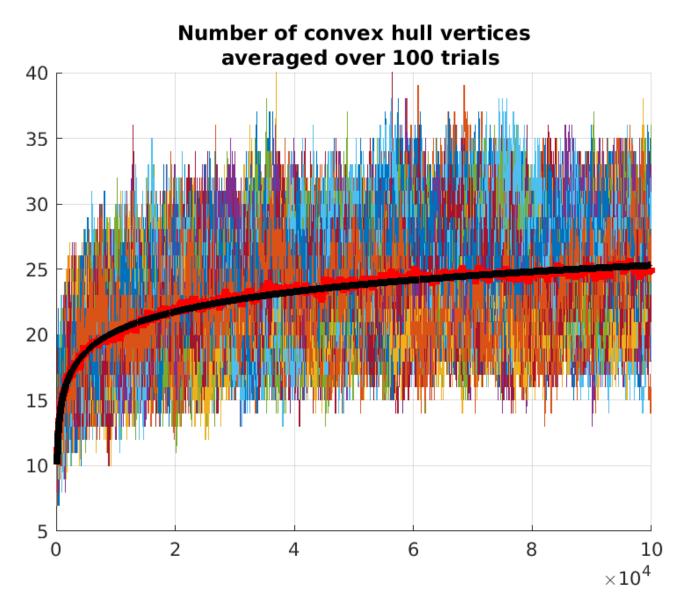


If convex hull has "sharp angles" some vertices have larger than average harmonic measure, implies faster than trivial growth.

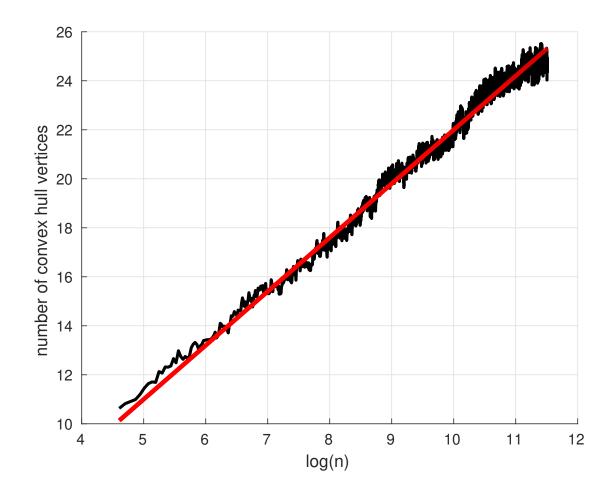


One way to have "sharp angles" is to have few vertices: if the convex hull boundary has few vertices, some of the angles should be large.

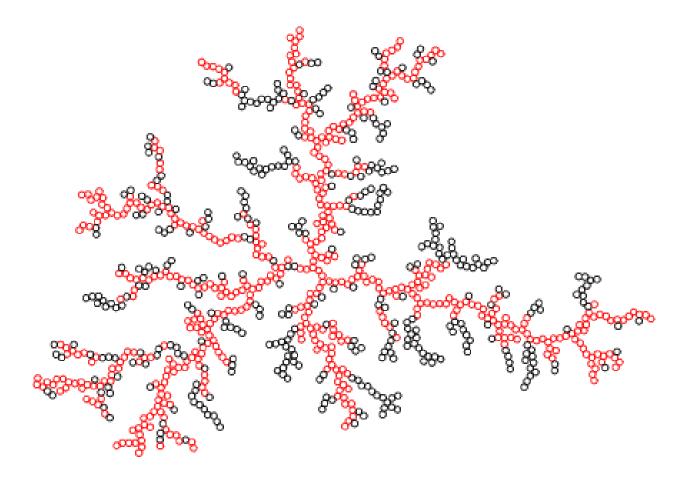




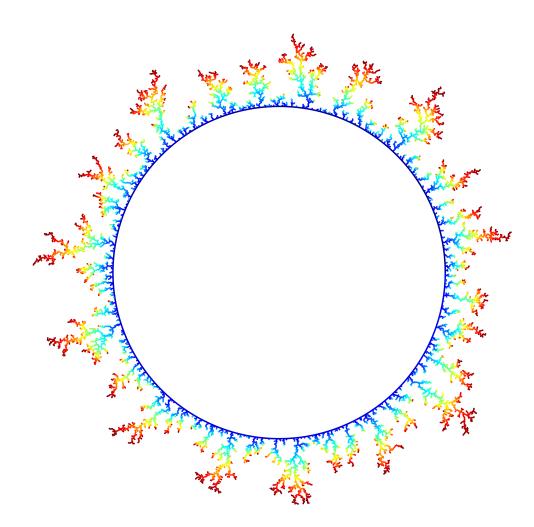
How many convex hull vertices are there at time n?



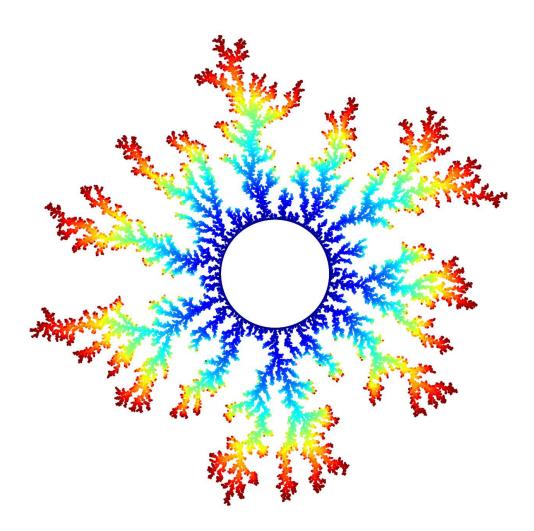
Number of convex hull vertices, averaged over 100 trials. Plotted versus  $\log(n)$ , looks linear. Numerically,  $\approx (2.2) \log n$ .



Red disks where on convex hull boundary when added. Percentage probably tends to zero, but how fast?

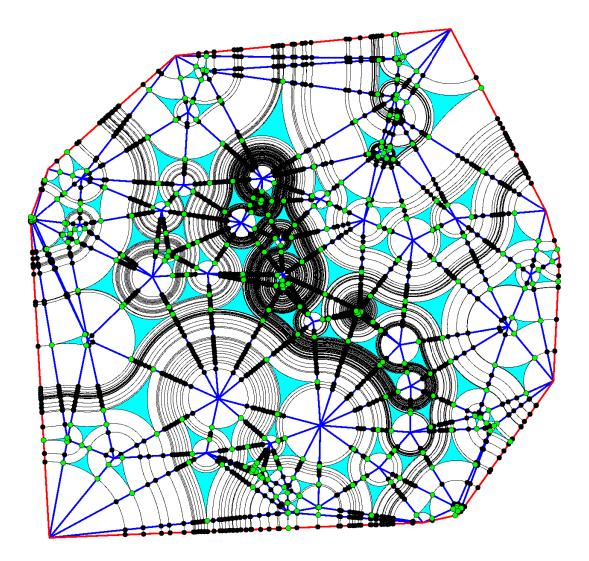


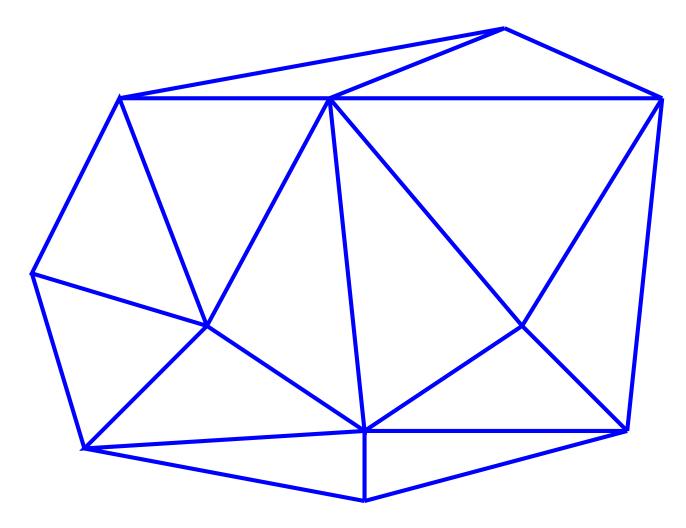
#### Prove that disk evolves into a "non-disk"



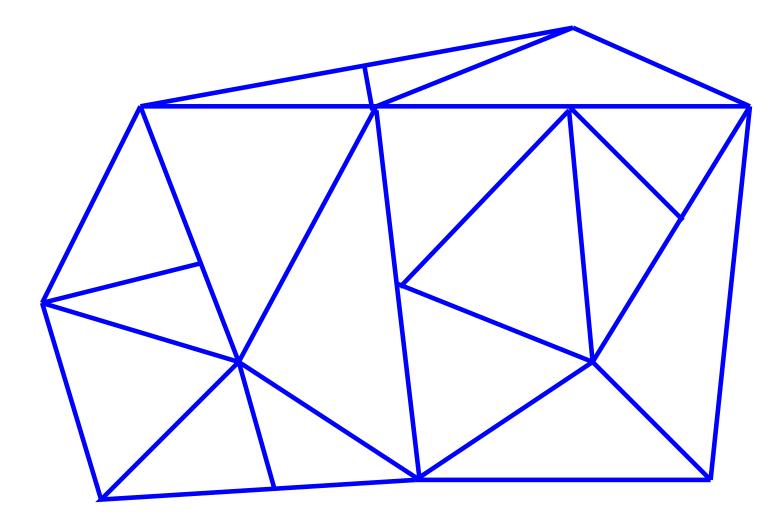
Thanks for listening. Questions?

# PART VI: TRIANGULATION FLOWS

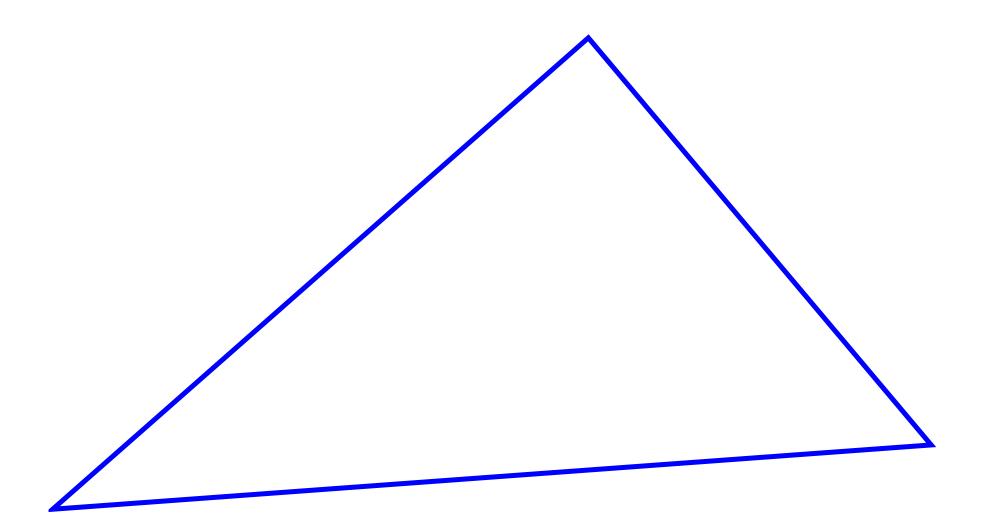




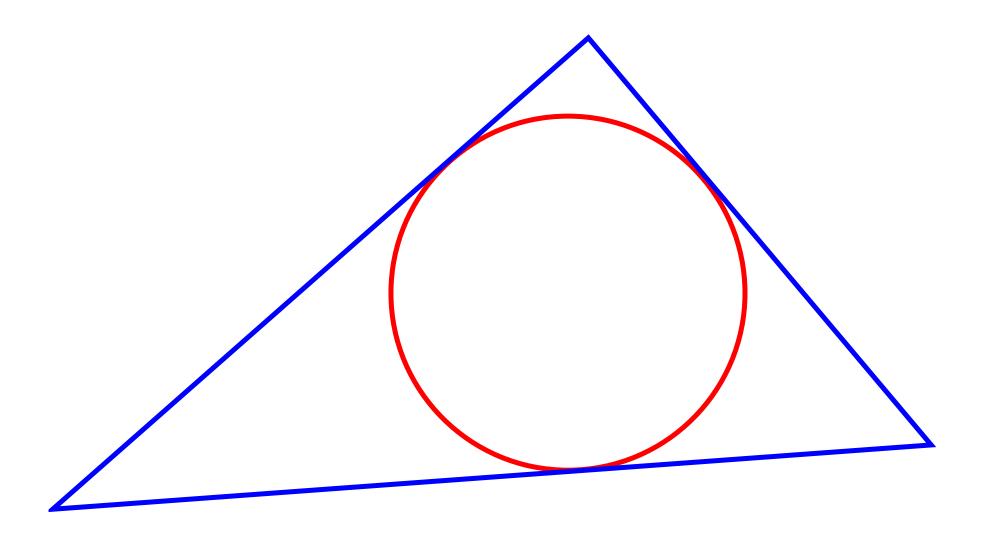
A triangulation: overlapping edges agree.



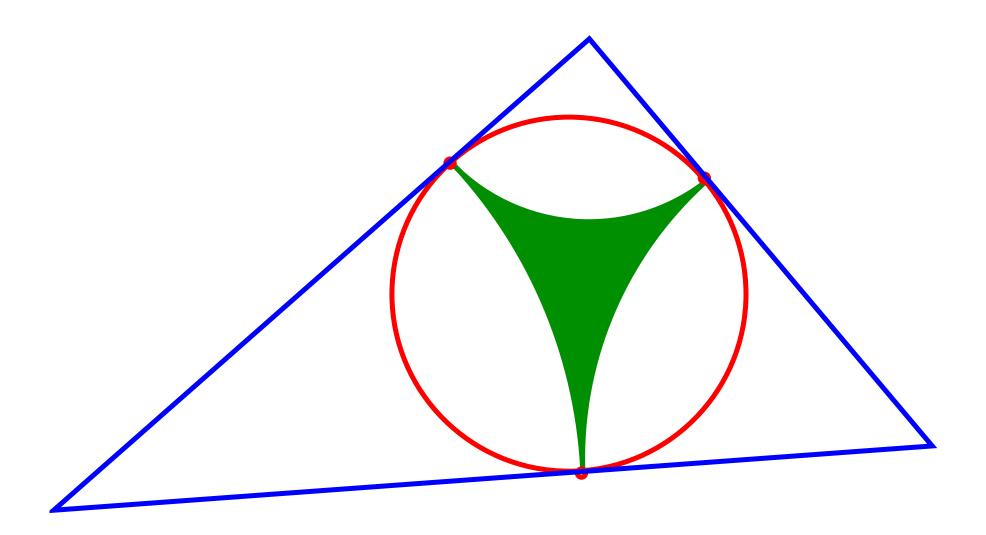
A dissection.



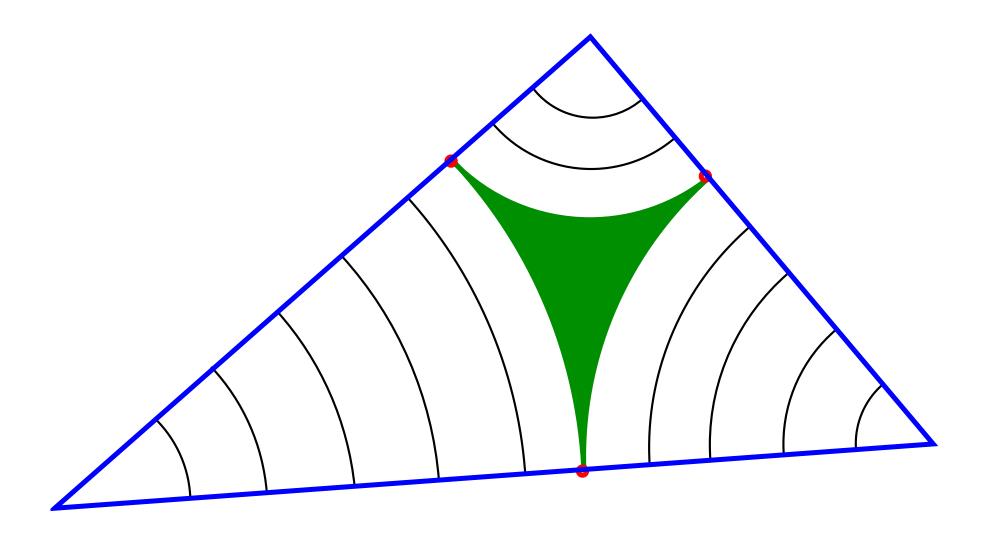
A triangle.



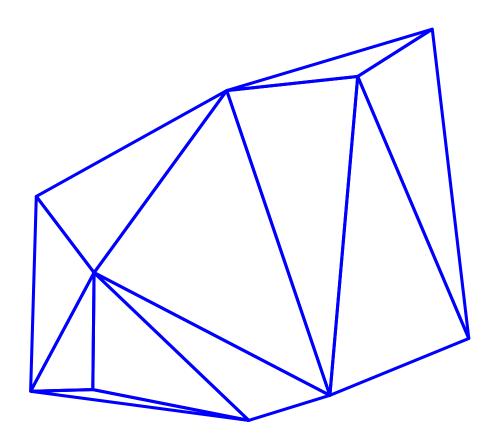
Its in-circle.



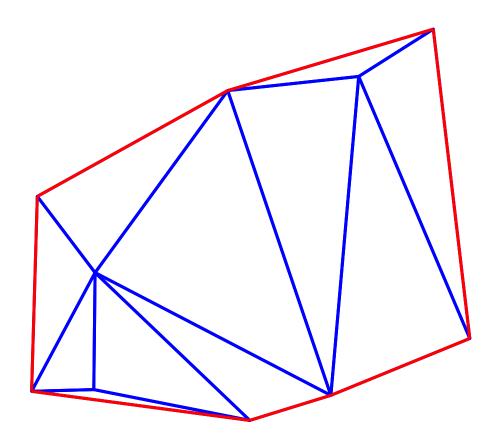
The central region and three sectors (thin version).



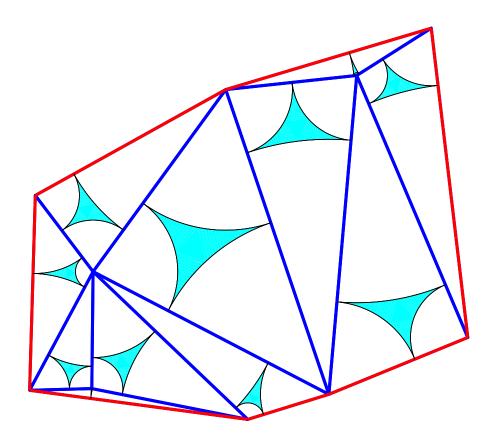
The three sectors are foliated by circular arcs. Defines flow on a triangulation that stops at boundary or cusp point.



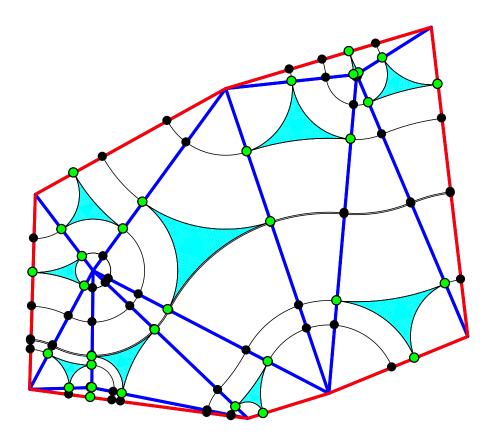
# Delaunay triangulation of 10 random points,



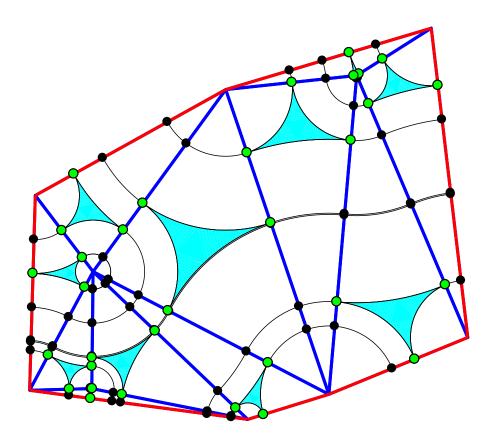
# The boundary of the triangulation



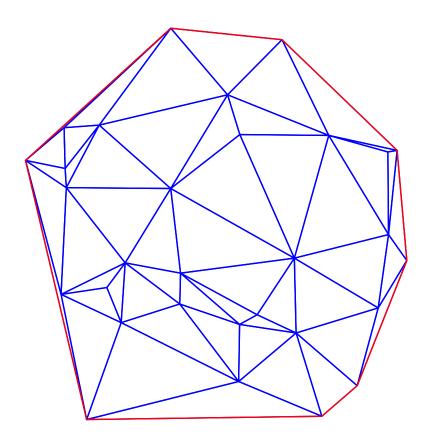
The central regions.



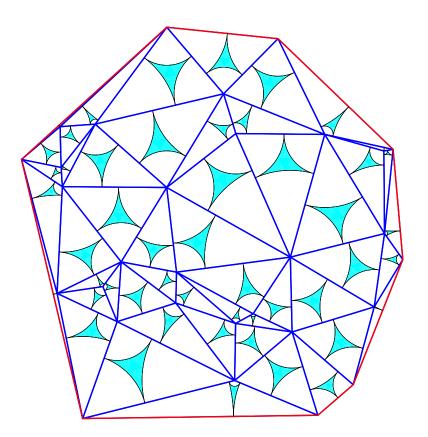
## Propagation lines starting at all cusp points.



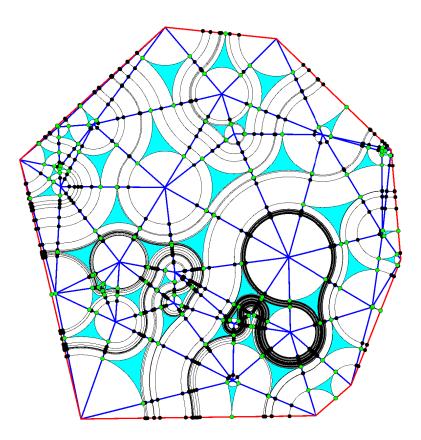
Propagation lines identify boundary points; induces tree. Discontinuous, but piecewise length preserving.



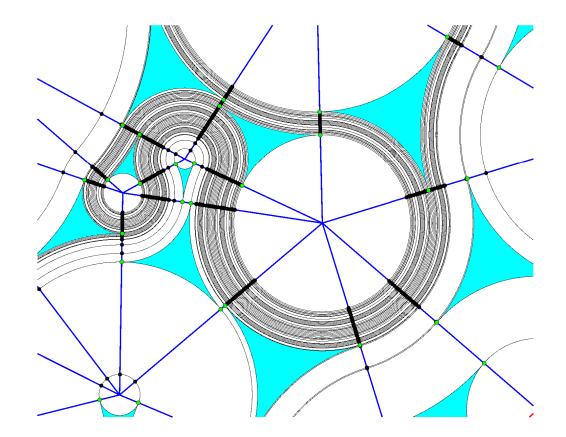
Delaunay triangulation of 30 random points in disk.



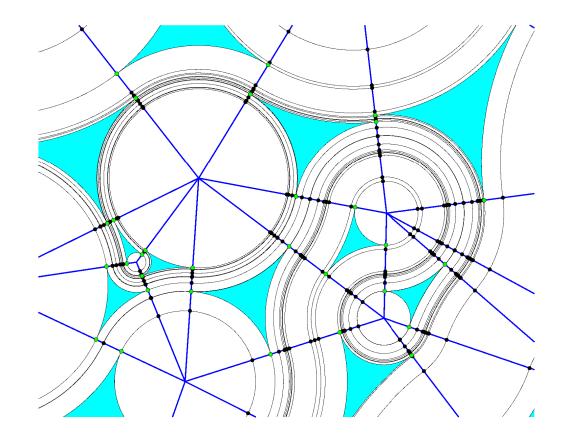
The central regions.



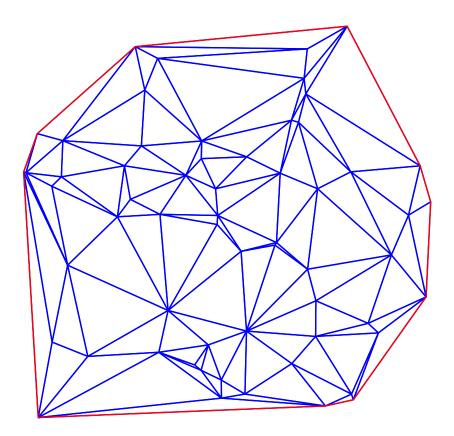
# Propagation lines starting at all cusp points.



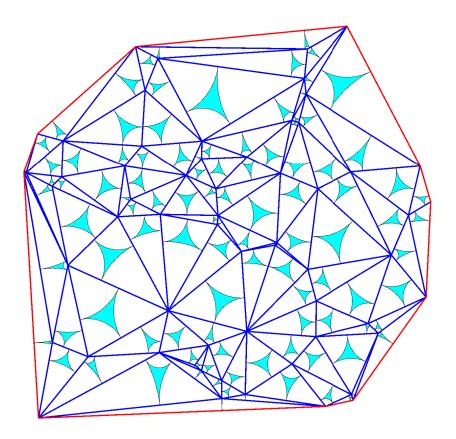
Enlargement 1.



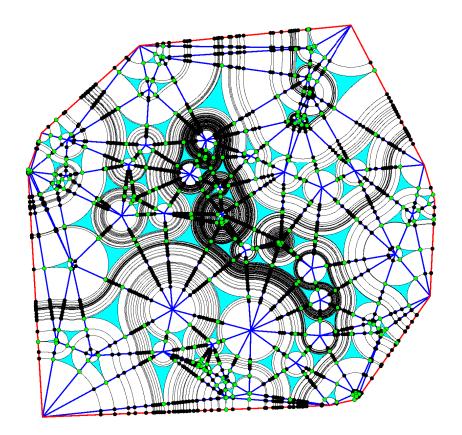
Enlargement 2.



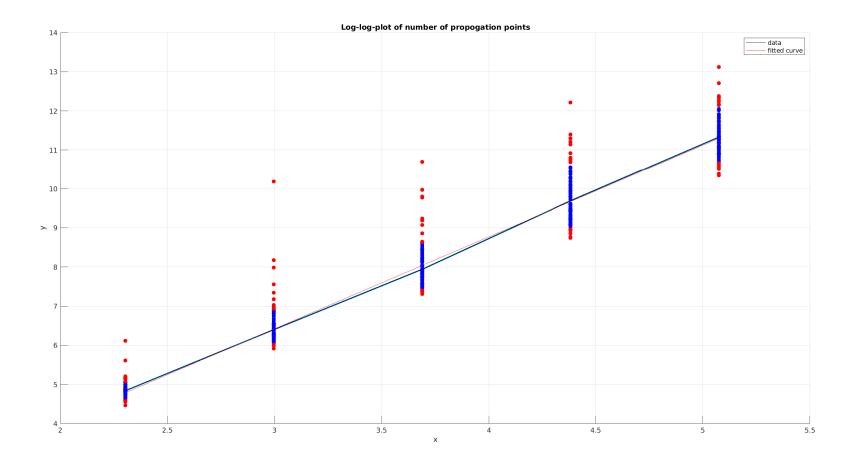
60 points



The central regions.



# Propagation lines starting at all cusp points.

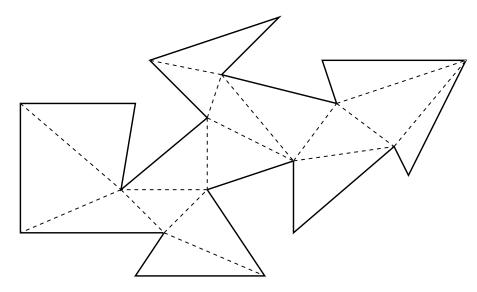


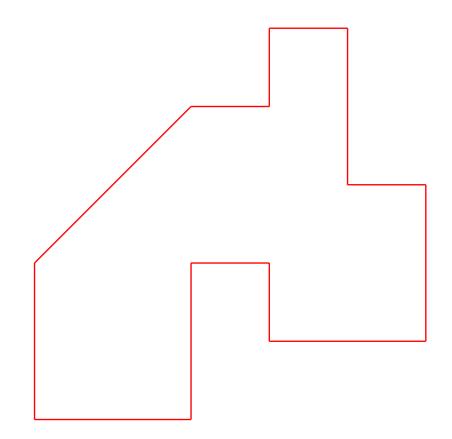
Log-log plot of number points created versus n. Slope  $\approx 2.5$ 

**Theorem:** For a triangulation of a simple polygon by diagonal, at most  $O(n^2)$  points are created.

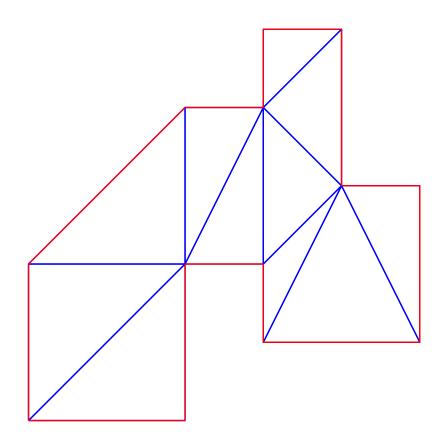
**Proof:** In this case, the triangles form the vertices of a tree where adjacency means sharing an edge.

Since a flow line never re-enters a triangle, it visits at most n triangles, so at most  $O(n^2)$  points are generated.

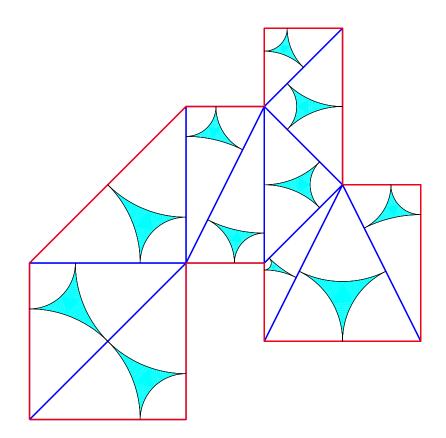




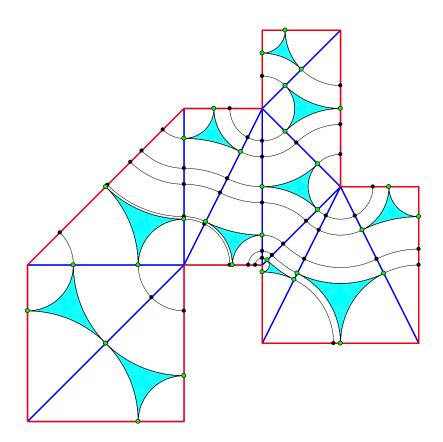
A simple polygon



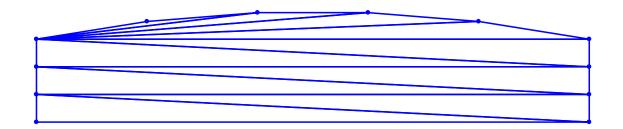
### A triangulation of the polygon using diagonals.



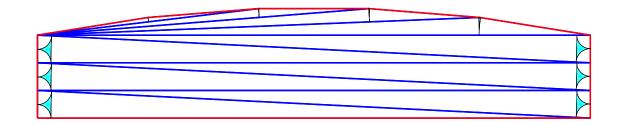
The central cusp regions.



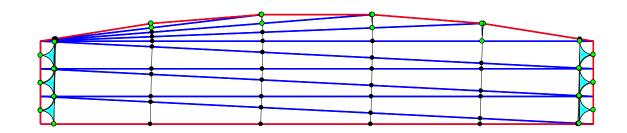
No flow line returns to a triangle, so each path creates at most n new points, for a total of  $O(n^2)$ .



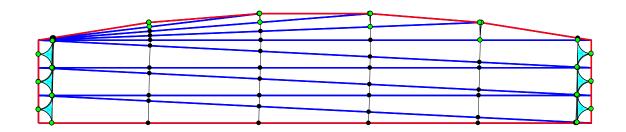
### The $n^2$ is sharp.



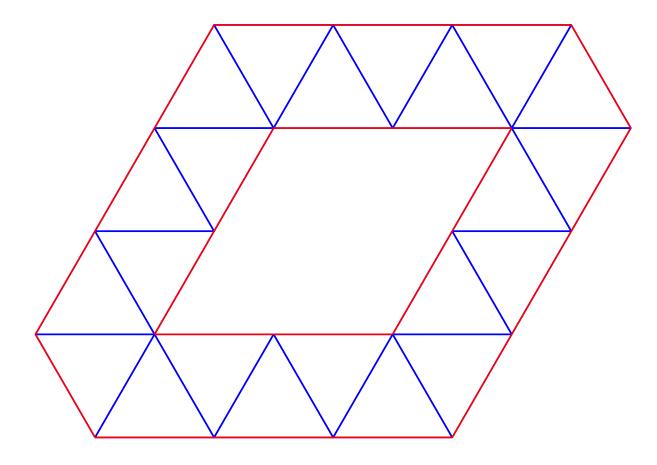
The  $n^2$  is sharp.



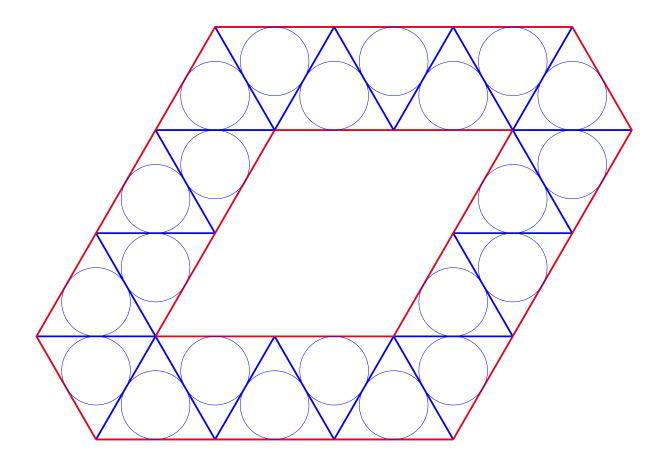
Each point "along top" creates a path that hits  $\simeq n$  triangles.



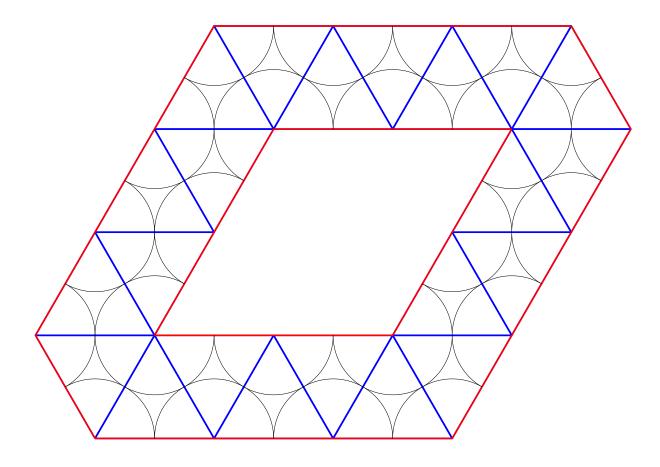
The  $n^2$  is stable under small perturbations of vertices.



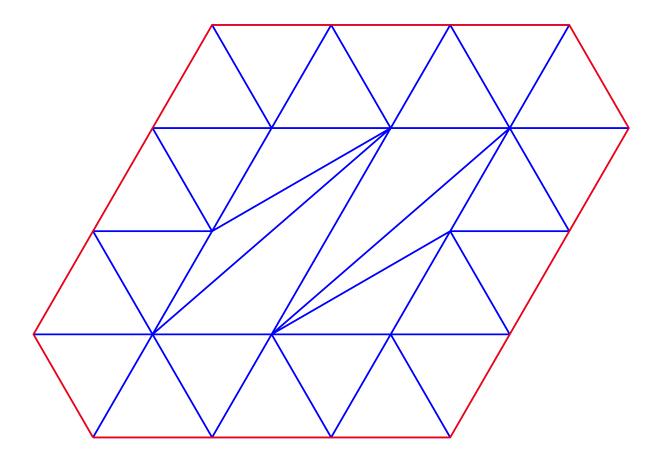
A ring of equilateral triangles.



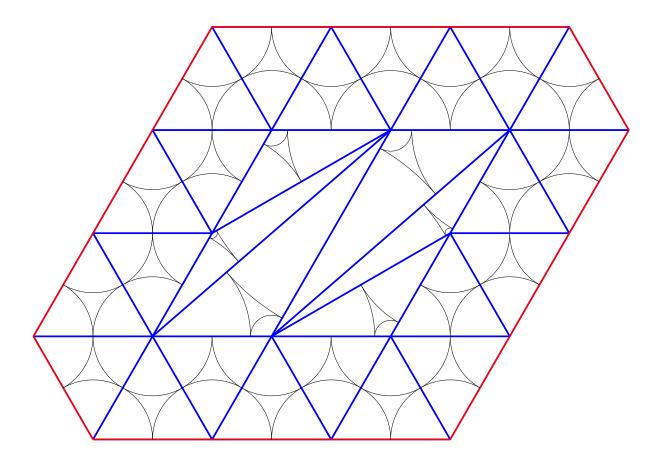
The in-circles are tangent.



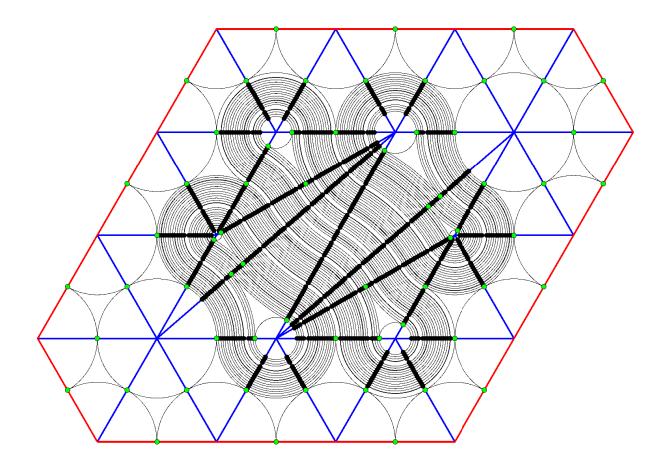
The the cusps touch; for a closed flow line.



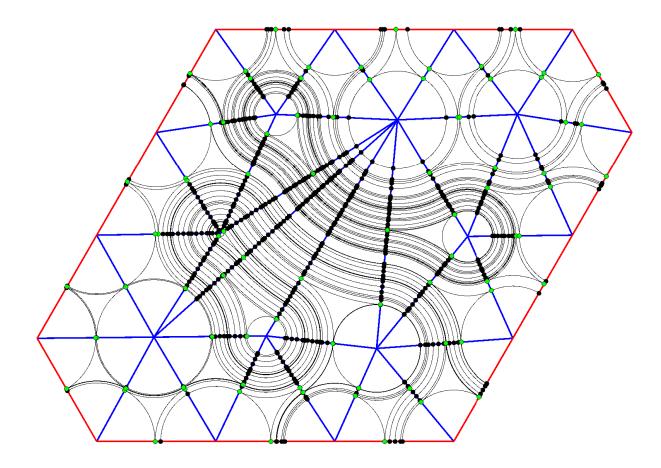
No matter how we triangulate interior, flow lines never exit.



No matter how we triangulate interior, flow lines never exit.



No matter how we triangulate interior, flow lines never exit.



Randomly perturb some vertices; flow lines "leak" to the boundary.

The triangulation flow arises from applications.

A non-obtuse triangulation (NOT) has angles  $\leq \pi/2$ .

NOTs important in applications, give better numerical results.

E.g., Vavasis showed that matrices arising from finite element method for a certain PDE have conditions numbers that grow exponentially (in number of triangles) for general triangulations, but only linearly for non-obtuse triangulations.

# **NOT Theorem, (B. 2016)** Any *n*-triangulation has a non-obtuse refinement with $O(n^{2.5})$ elements.

**NOT Theorem, (B. 2016)** Any *n*-triangulation has a non-obtuse refinement with  $O(n^{2.5})$  elements.

First polynomial time bound. 2.5 sharp?

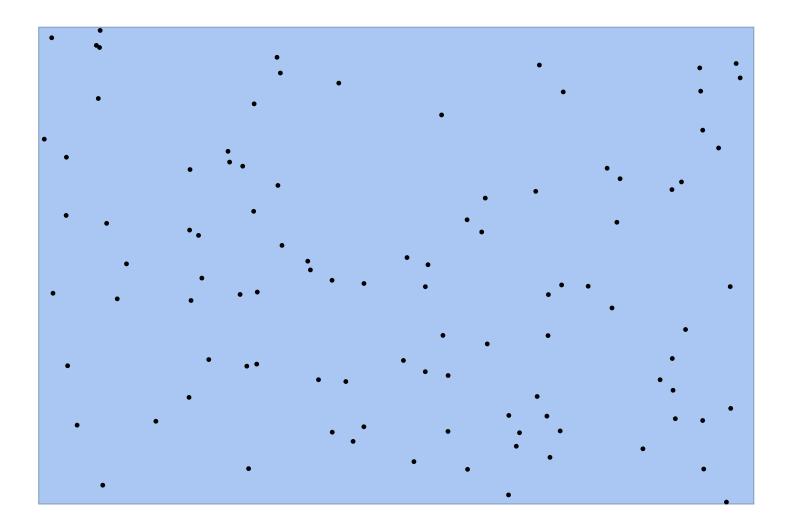
Crossing points of triangle flow give triangulation vertices.

If too many crossings, perturb flow to create closed loops.

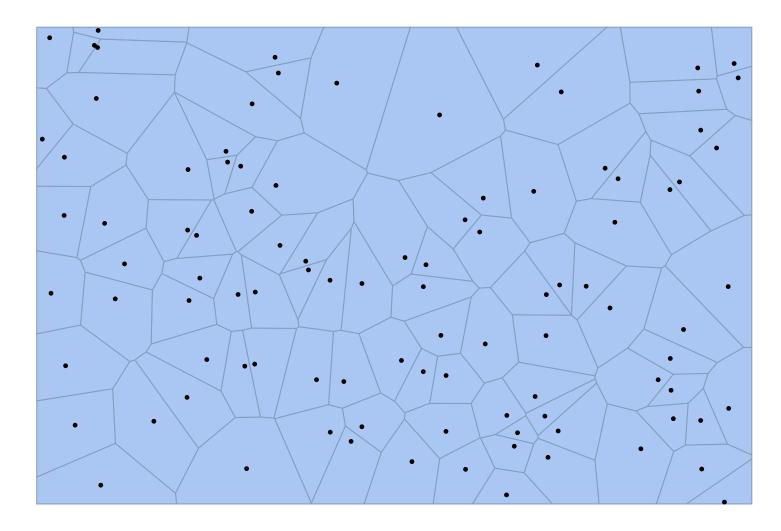
Reminiscent of Pugh's closing lemma in dynamics.

Gives  $O(n^{2.5})$  in worst case. What is average behavior?

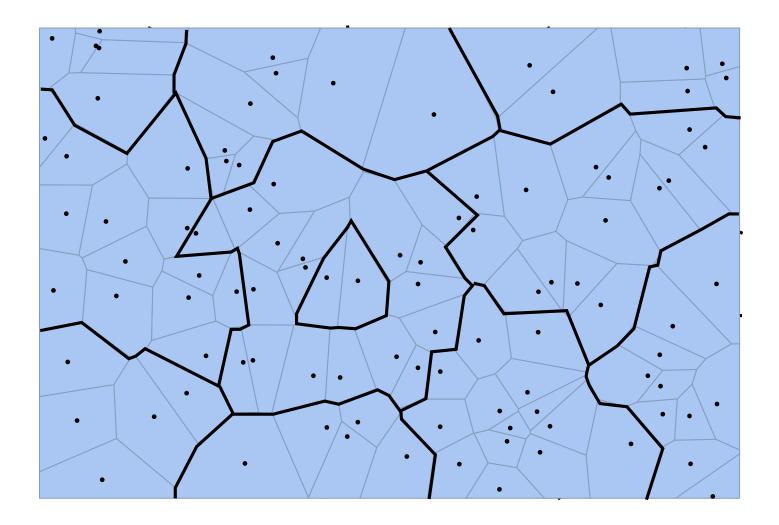
## THANKS FOR LISTENING QUESTIONS?



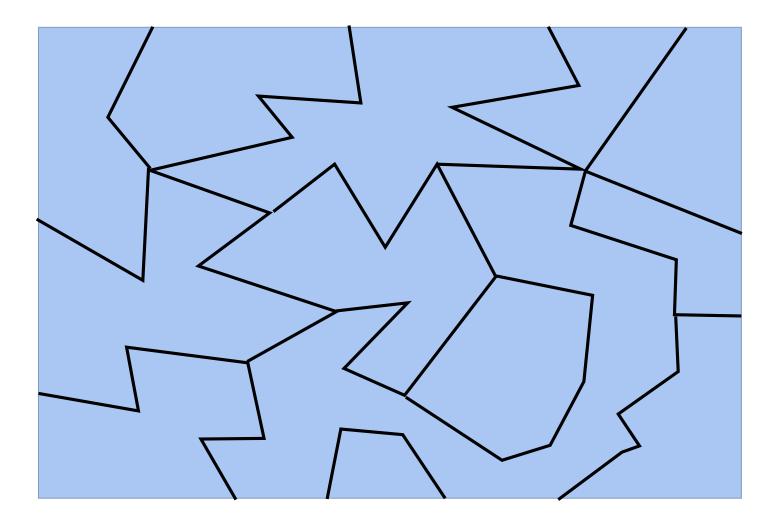
An application of the NOT theorem Consider a finite set of points in the plane.



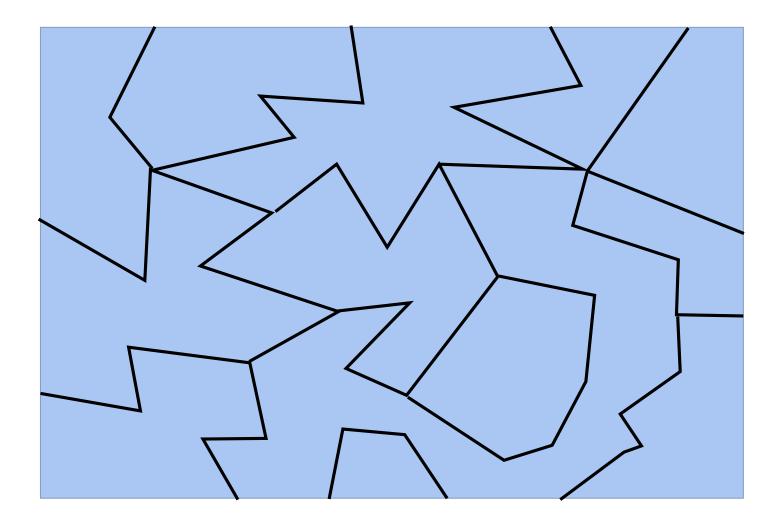
Voronoi cells (think of cell phone connecting to closest tower).



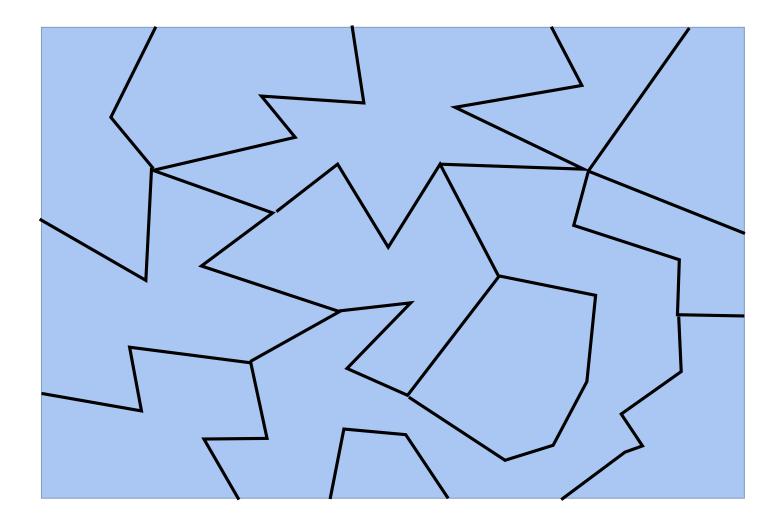
If region boundaries conform to cell boundaries, then a phone always connects to a tower in the same region.



Given countries, can we place towers so this happens? Do a polynomial number of towers suffice?



Given countries, can we place towers so this happens? Do a polynomial number of towers suffice? Yes (B 2016)



Proof: It's easy to place points explicitly if regions are all non-obtuse triangles. In general, triangulate the regions, then non-obtusely refine the triangulation.

Theorem (Kesten): diam(DLA(n)) =  $O(n^{2/3})$ .

Equivalent: DLA takes  $\gtrsim m^{3/2}$  steps to exit ball of radius m.

Suppose current radius is m. How long to reach 2m?

Sketch (following Lawler). Suppose  $\beta m^{3/2}$  disks suffice to exit,  $\beta > 0$ .

Then cluster contains a chain of lattice points  $\mathbf{z} = \{z_1, \ldots, z_k\}$  so that

$$|z_{1}| < m/2, \qquad |z_{k}| > m,$$

$$|z_{j} - z_{j+1}| \le 4, j = 1, \dots, k, \qquad m/4 \le k \le \beta m^{3/2}.$$

Let  $W_m(\mathbf{z})$  be all clusters associated to a chain  $\mathbf{z}$ . Let  $W_m = \bigcup_{\mathbf{z}} W_m(\mathbf{z})$ .

At most  $O(m^2 80^k)$  chains:  $O(m^2)$  starting points and 80 choices per step.

Claim:  $\operatorname{Prob}(W_m(\mathbf{z})) \leq (C\beta)^k$ .

Let  $W_m(\mathbf{z})$  be all clusters associated to a chain  $\mathbf{z}$ . Let  $W_m = \bigcup_{\mathbf{z}} W_m(\mathbf{z})$ .

At most  $O(m^2 80^k)$  chains:  $O(m^2)$  starting points and 80 choices per step.

Claim:  $\operatorname{Prob}(W_m(\mathbf{z})) \leq (C\beta)^k$ .

Assuming claim, Kesten's theorem follows: if  $\beta$  small,

$$\sum_{m} \operatorname{Prob}(W_m) \le \sum_{m} \sum_{\mathbf{z}} \operatorname{Prob}(W_m(\mathbf{z})) \le C \sum_{m} m^2 80^m C^m \beta^m < \infty.$$

By Borel-Cantelli a.s.  $W_m$  occurs only finitely often.

Eventually, exit time from radius m is  $> \beta m^{3/2}$ . QED

#### **Proof of claim:**

Given a chain  $\mathbf{z} = z_1, \ldots, z_m$ , let  $D_j$  is disk covering  $z_j$ .

How long do we wait between adding  $D_j$  and  $D_{j+1}$ ?

Probability that next disk lands less than distance 4 of  $D_j$  is  $\leq m^{-1/2}$ .

Why?

Well known result in conformal mapping.

**Beurling's thm:** If  $\Omega = \mathbb{C} \setminus K$ , K compact and connected,  $x \in K$ ,

$$\omega(\infty, D(x, 1) \cap K, \Omega) \le \frac{C}{\sqrt{\operatorname{diam}(K)}}.$$

Suppose  $X_k$  is event " $D_{j+1}$  is added to cluster  $\geq k$  steps after  $D_j$ ".

Then  $\operatorname{Prob}(X_k) \leq (1-p)^t$  where  $p \leq m^{1/2}$ .

Fact from Probability: If  $X_1, \ldots, X_n$  are independent geometric random variables with parameter p, then

$$\operatorname{Prob}(\sum_{k=1}^{n} X_k < \frac{\beta n}{p}) \le (2e^2\beta)^n.$$

n = number of points in chain  $\in [m, m^{3/2}]$  $\sum_{k=1}^{n} X_k =$  time to cover all points in chain  $\frac{\beta n}{p} \ge \frac{\beta m}{m^{-1/2}} = \beta m^{3/2}$ 

Prob(chain is covered in  $\leq \beta m^{3/2}$  steps)  $\leq (2e^2\beta)^n$ .

 $\Rightarrow$  Kesten's theorem.