From Random Polygon to Ellipse

A Snapshot of Computational Science and Engineering

Charlie Van Loan
Department of Computer Science
Cornell University

Outline

1. Tell the story behind my favorite research paper. It began as a "CS 1" assignment!
2. The problem is a metaphor for matrix-based computational science and engineering. Explain why.

But first, some context!

My Goal in CS 1

Develop a practical **intuition** about computer problem-solving and its role in science and engineering.

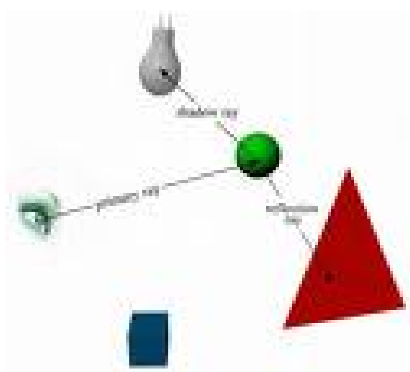
What do we mean by “Intuition“?

If intuition is a sense of direction, then computational intuition is a sense of computational direction.

The Five Computational Senses

1. A sense of geometry.
2. A sense of complexity.
3. A sense of randomness.
4. A sense of approximation.
5. A sense of error.

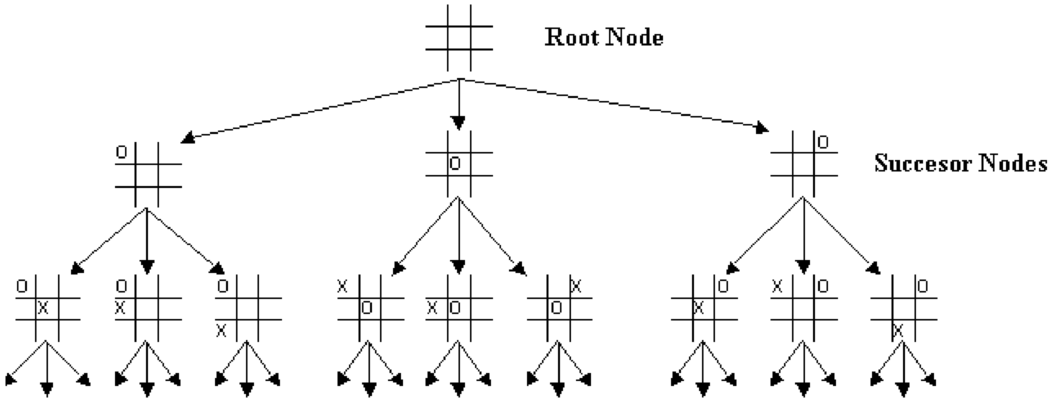
A Sense of Geometry



Ray Tracing

100 million triangles

A Sense of Complexity

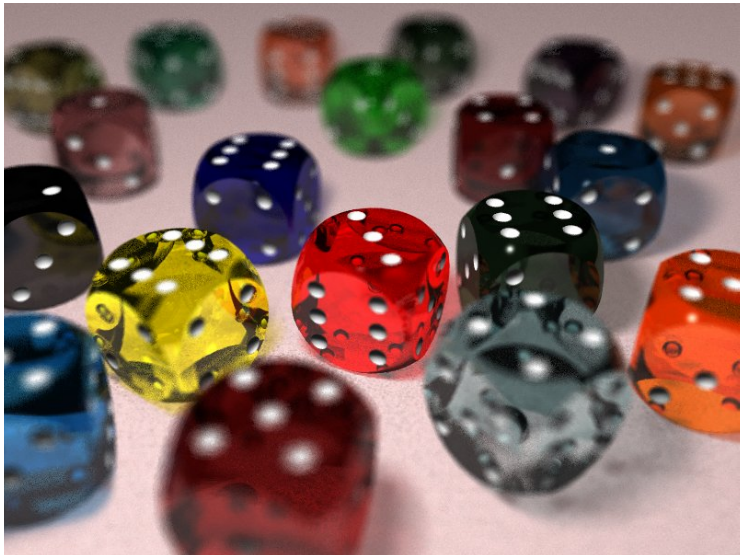


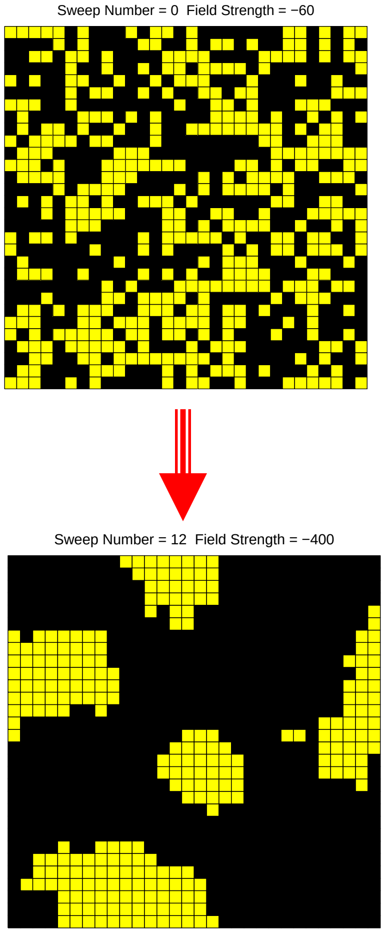
Search Trees

Design Space

Billions of Choices

A Sense of Probability and Statistics Via Simulation

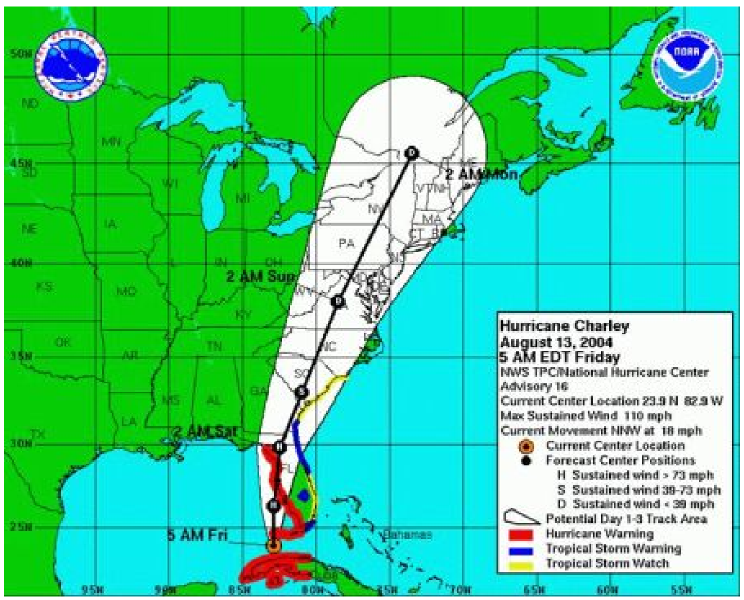




A Sense of Approximation, Error, and Noise

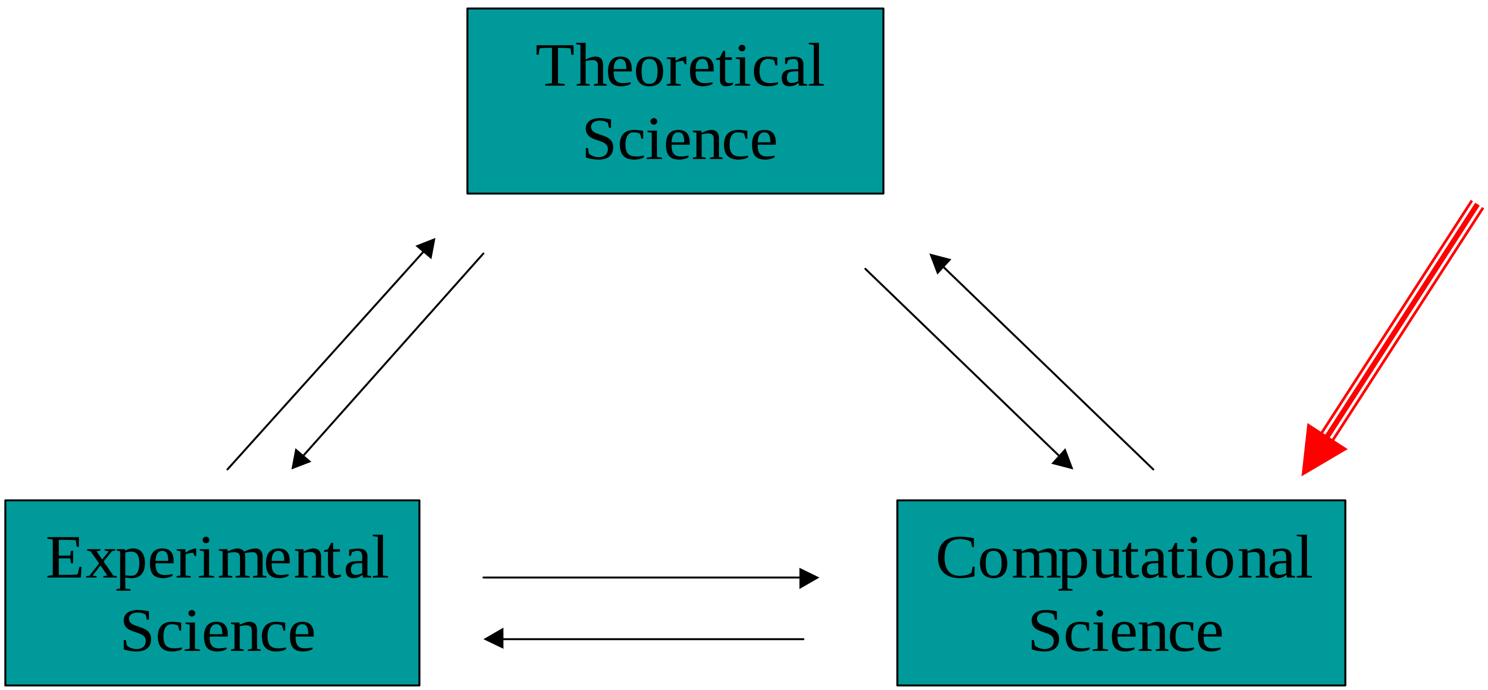
1/3 = .3333

Pi = 22/7



Captured Photo

Enter the Computer

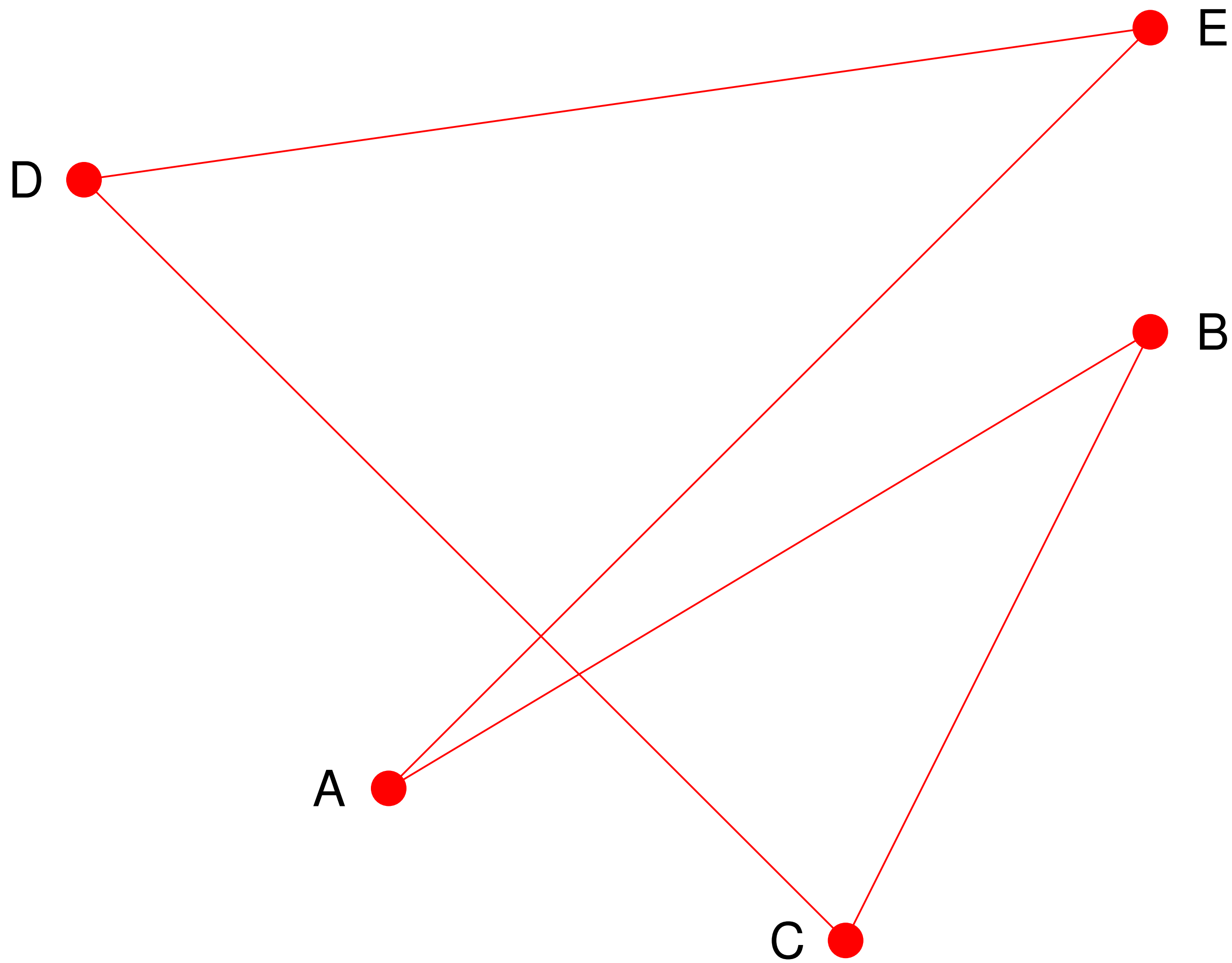


A Freshman Programming Problem

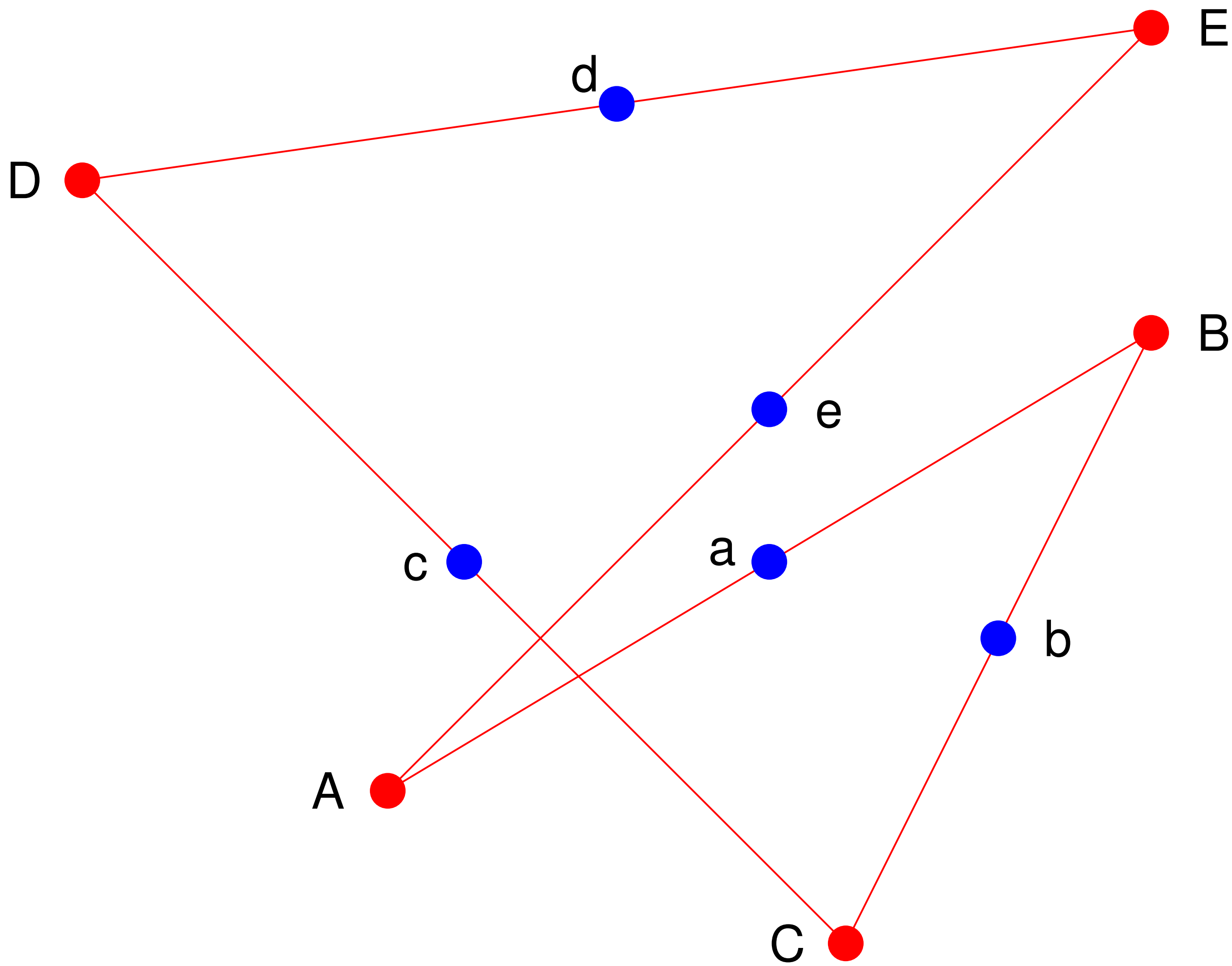
Given a polygon, generate a new polygon by connecting the midpoints of its sides.

Repeat the process many times and show what happens graphically.

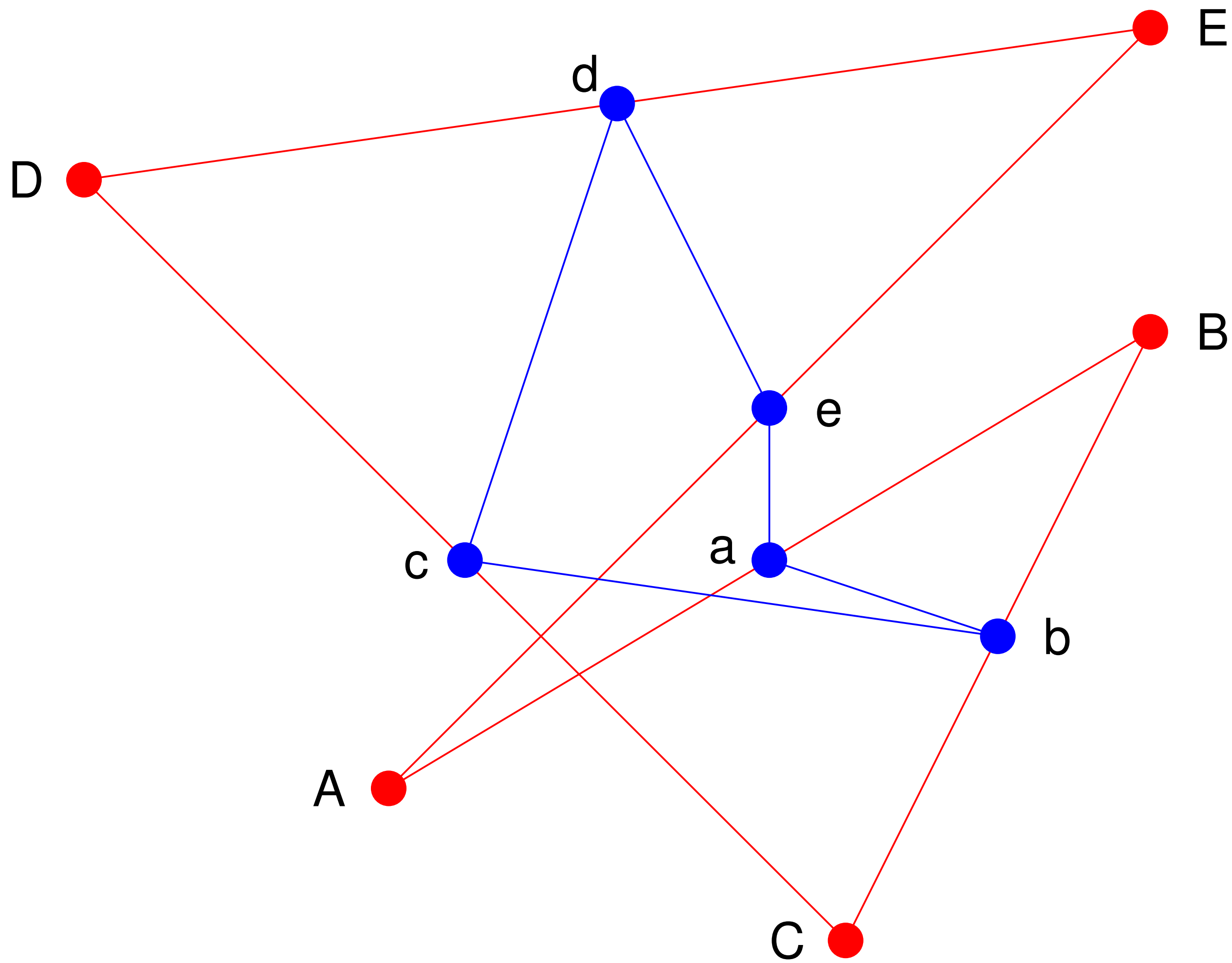
A Polygon..



The Midpoints...



The New Polygon...

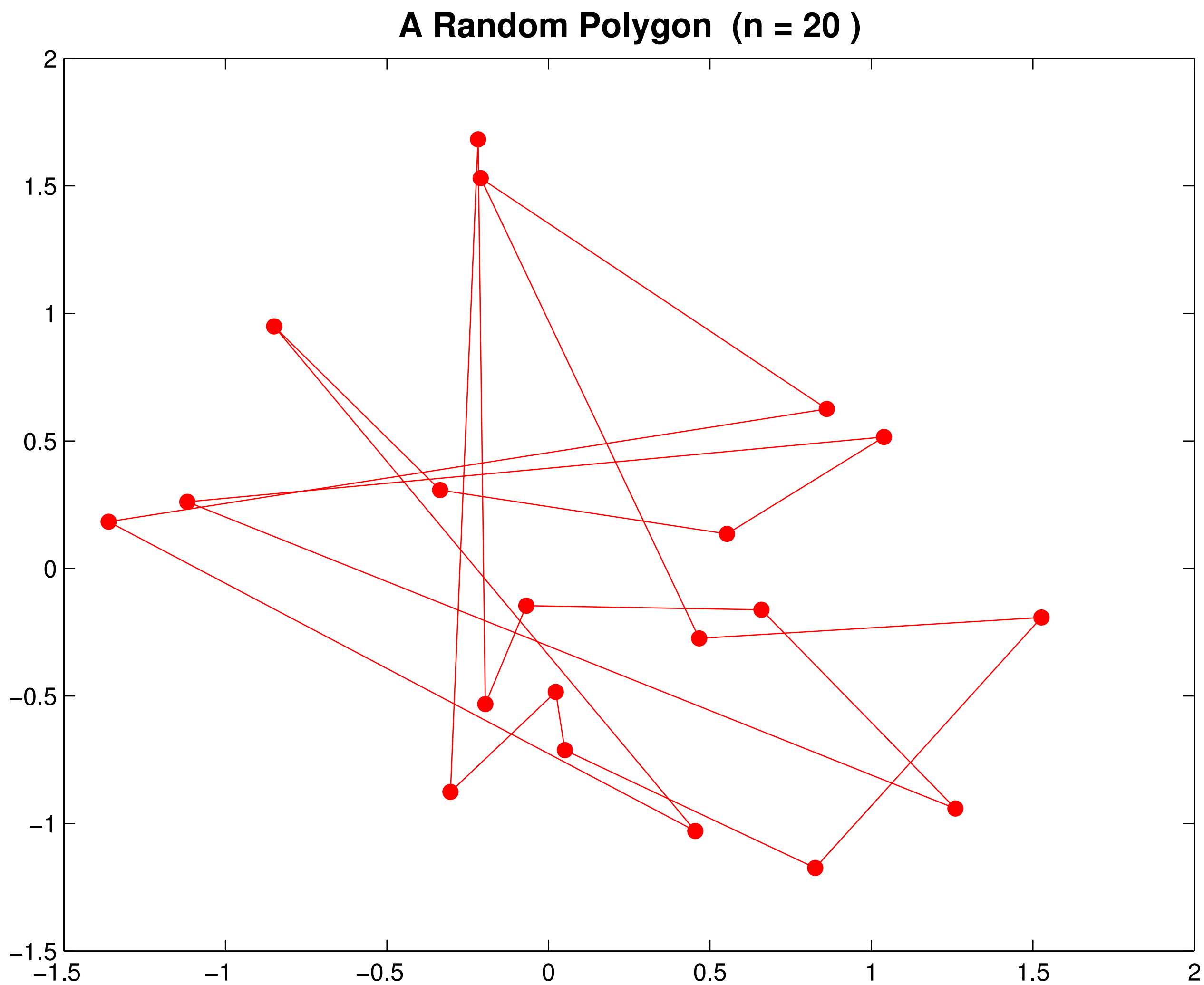


The Average of Two Polygons

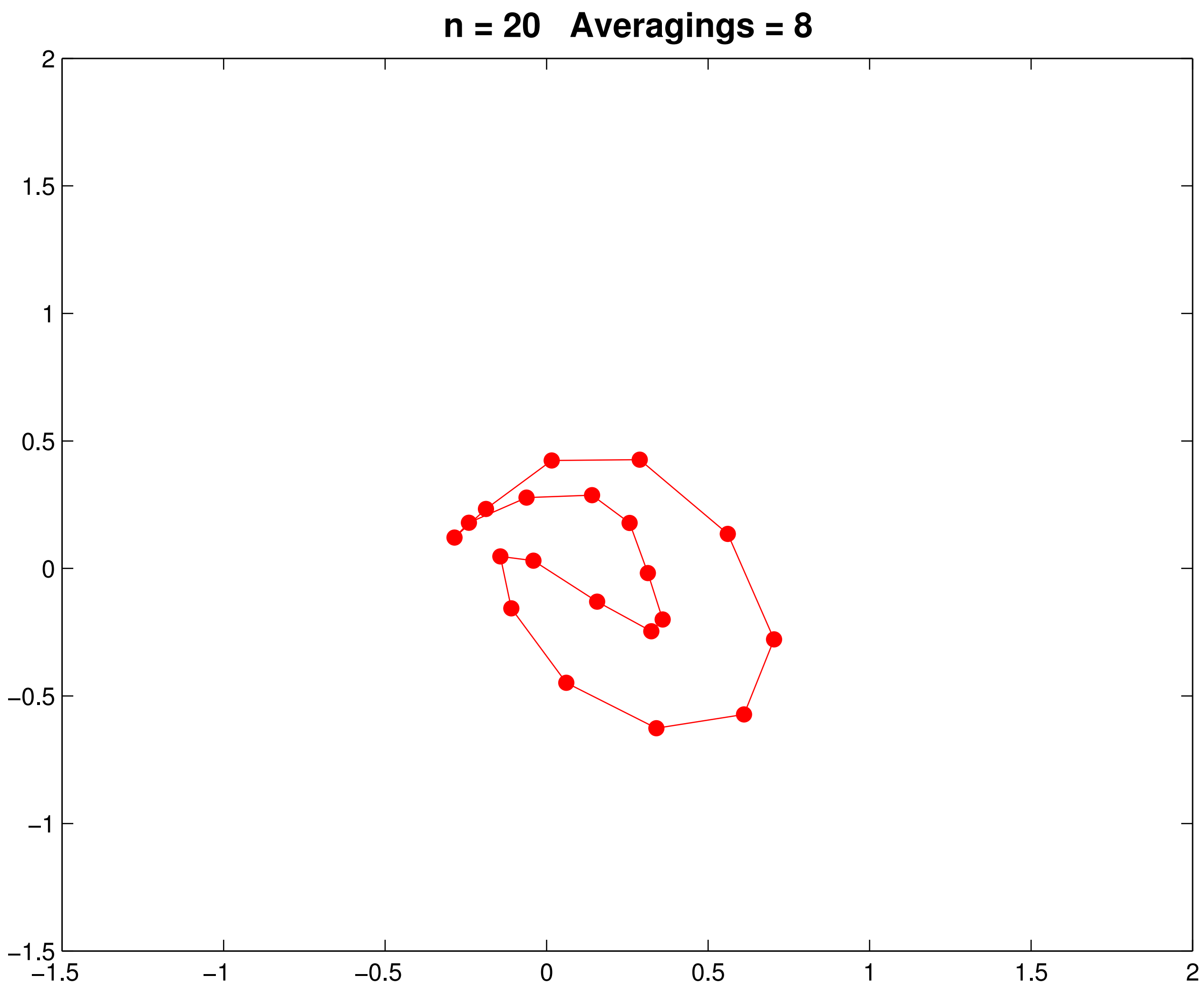
```
     x = [ 0.0  5.0   3.0  -2.0  5.0 ];
xShift = [ 5.0  3.0  -2.0   5.0  0.0 ];
  xNew = [ 2.5  4.0   0.5   1.5  2.5 ];

     y = [ 1.0  4.0   0.0   5.0  6.0 ];
yShift = [ 4.0  0.0   5.0   6.0  1.0 ];
  yNew = [ 2.5  2.0   2.5   5.5  3.5 ];
```

$$\mathcal{P}_{New} = \left(\mathcal{P} + \mathcal{P}_{Shift}\right)/2$$

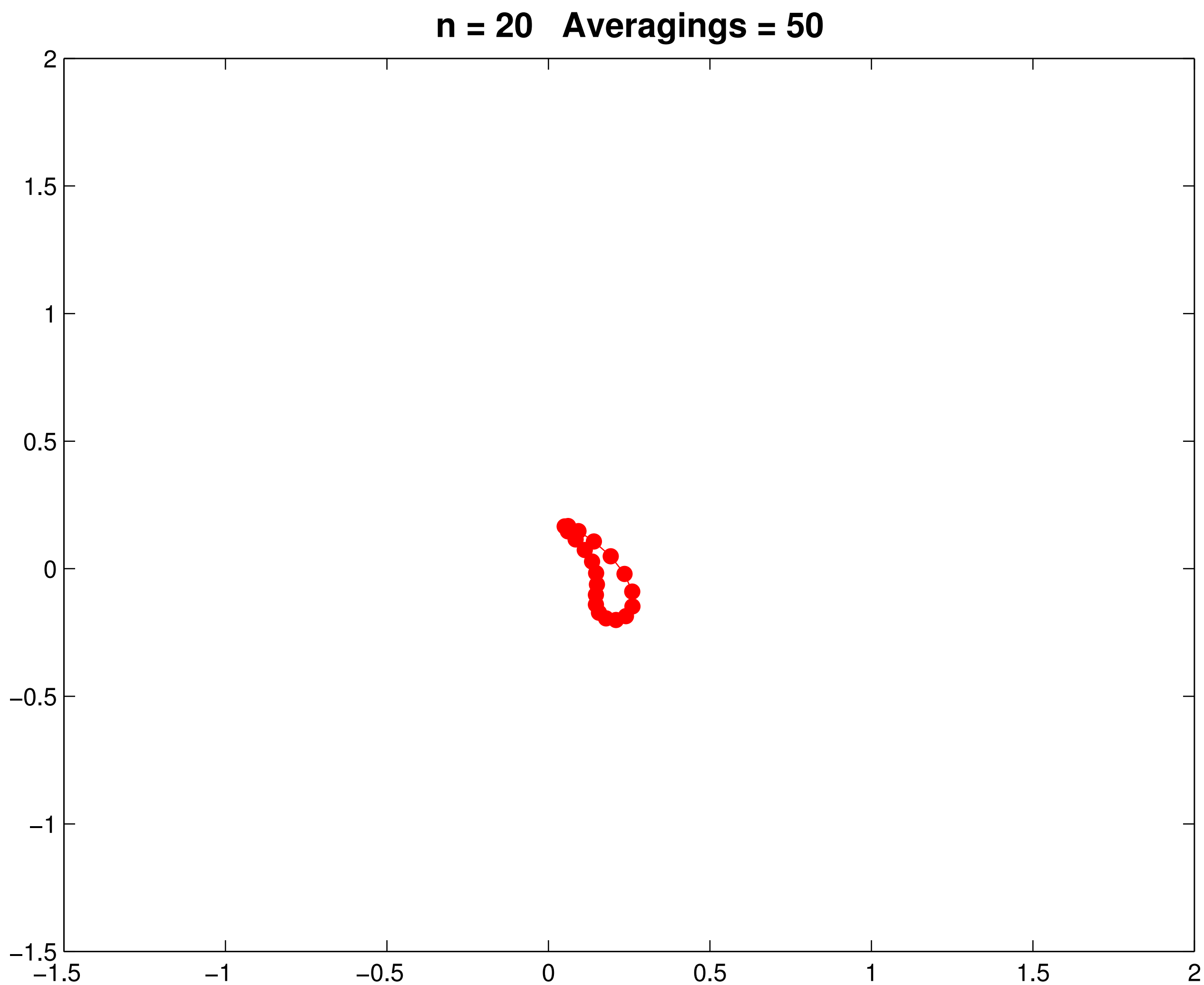
Solution 1



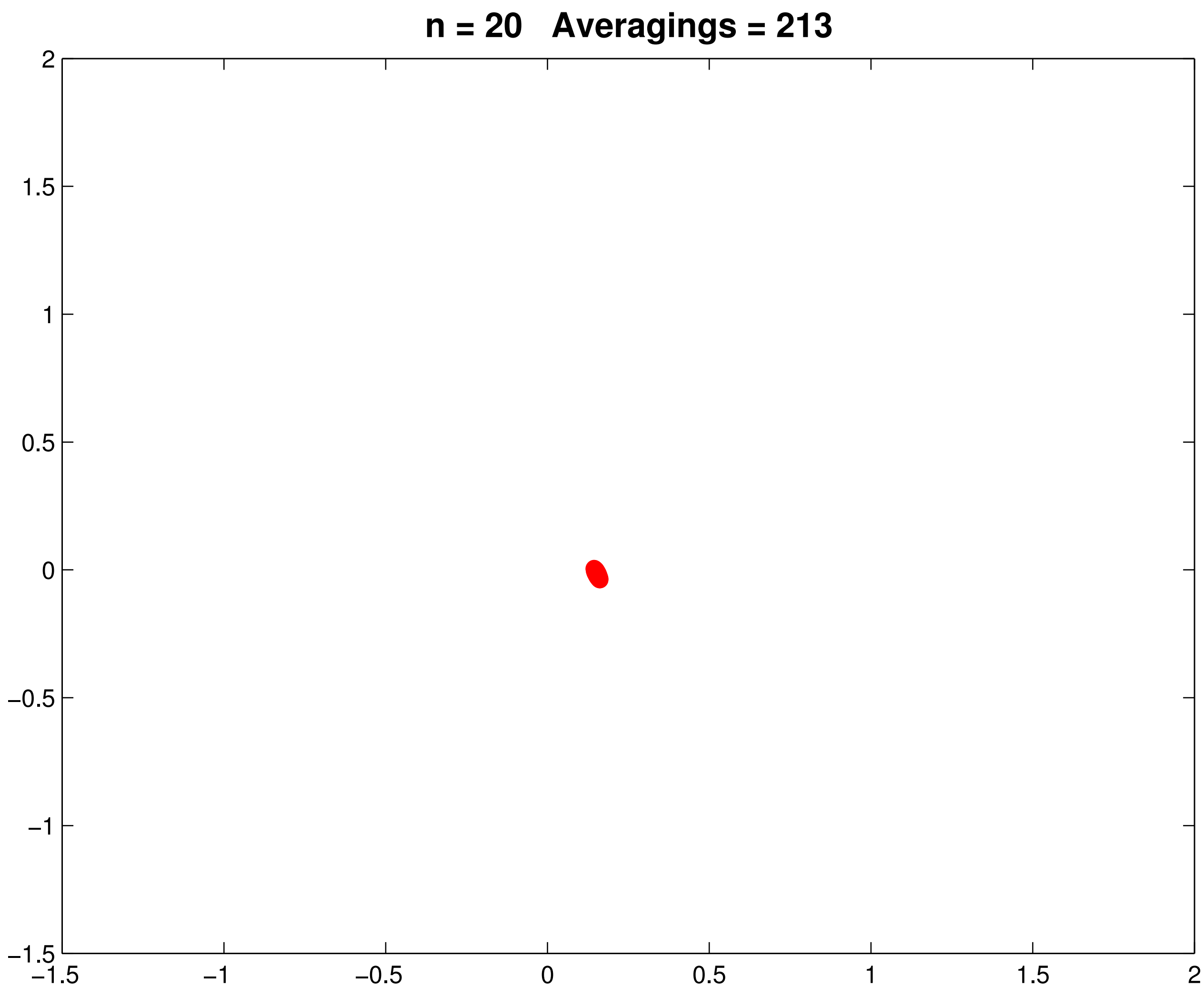
Solution 1



Solution 1



Solution 1

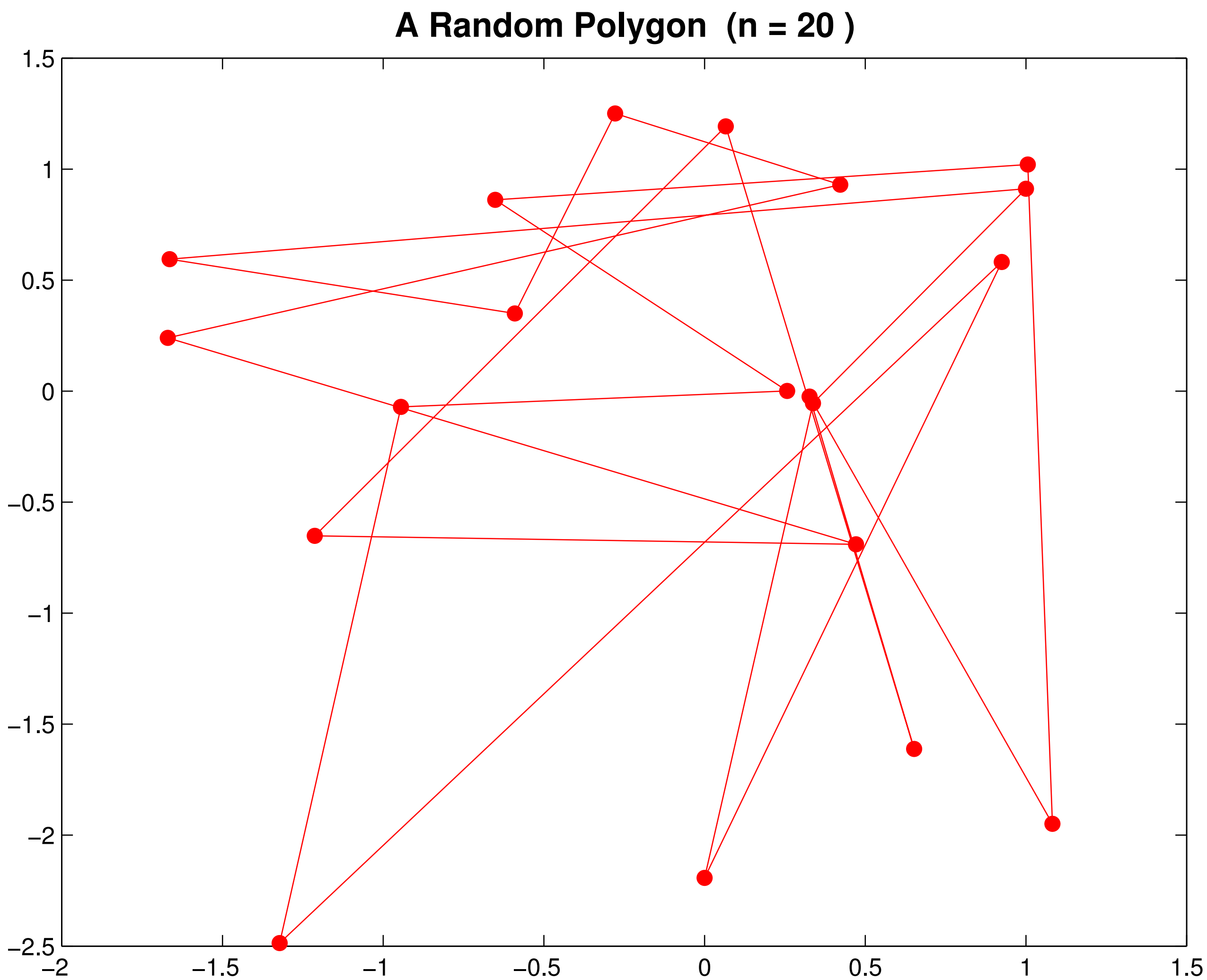


A Freshman Programming Problem (Revised)

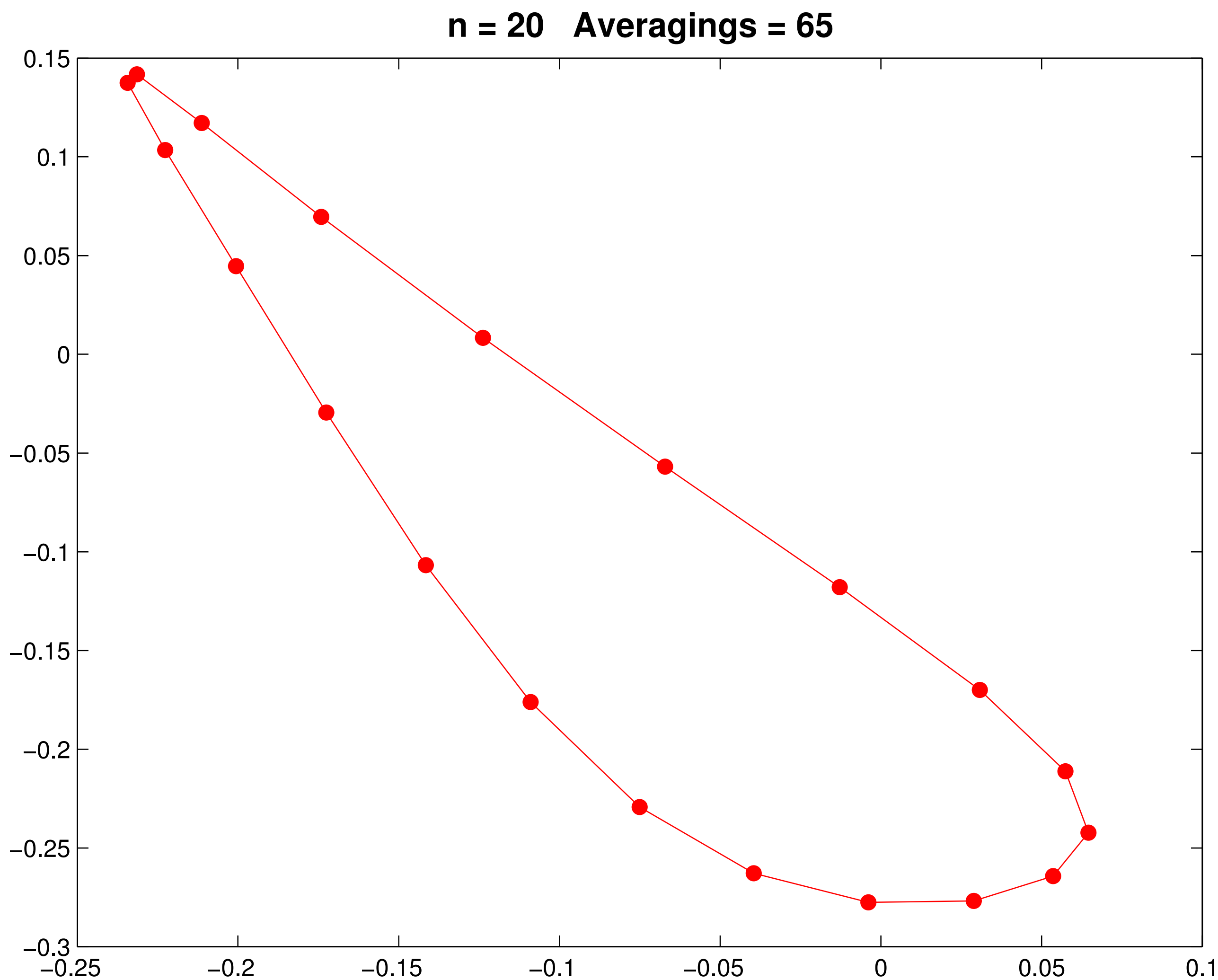
Given a polygon, generate a new polygon by connecting the midpoints of its sides.

Repeat the process many times and show what happens graphically. Use autoscaling.

Solution 2

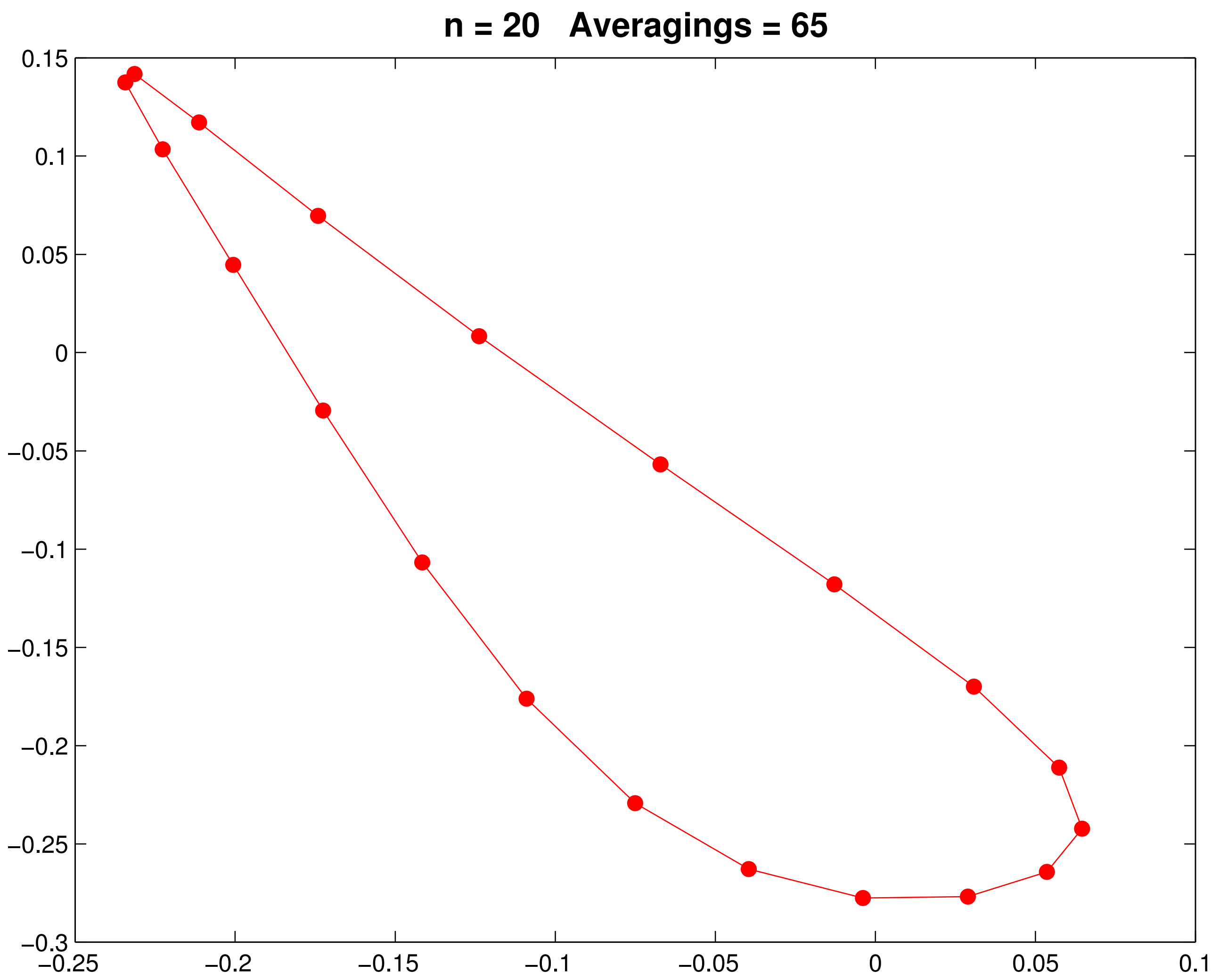


Solution 2



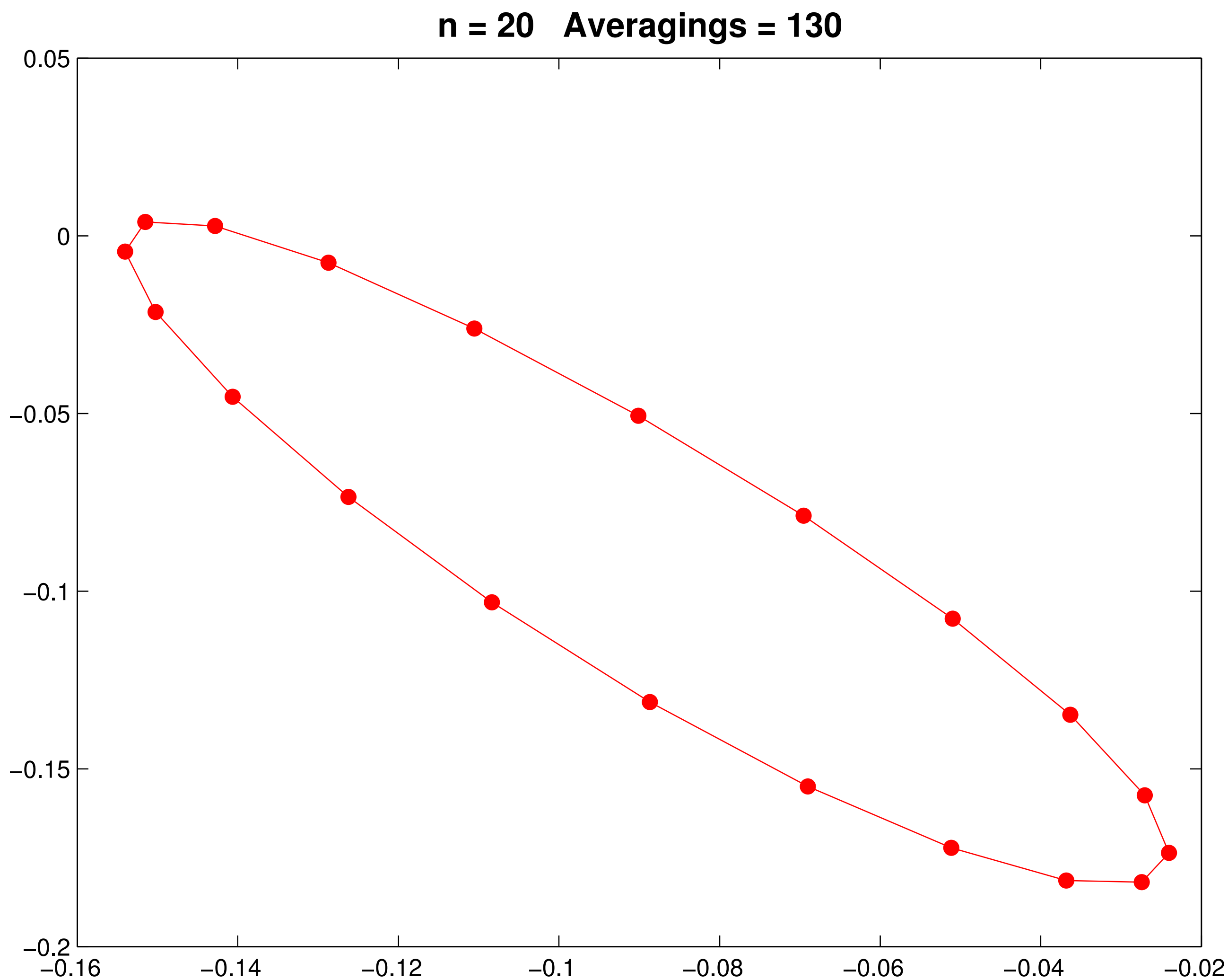
n = 20 Averagings = 65
0.15
0.1
0.05
0
−0.05
−0.1
−0.15
−0.2
−0.25
−0.3
−0.25
−0.2
−0.15
−0.1
−0.05
0
0.05
0.1

Solution 2



n = 20 Averagings = 65
0.15
0.1
0.05
0
−0.05
−0.1
−0.15
−0.2
−0.25
−0.3
−0.25
−0.2
−0.15
−0.1
−0.05
0
0.05
0.1

Solution 2



n = 20 Averagings = 130
0.05
0
−0.05
−0.1
−0.15
−0.2
−0.16
−0.14
−0.12
−0.1
−0.08
−0.06
−0.04
−0.02

A Senior-Level Scientific Computing Problem

Given a polygon with centroid (0,0), generate a a new polygon by connecting the midpoints of its sides. The x and y vertex vectors should always have unit 2-norm.

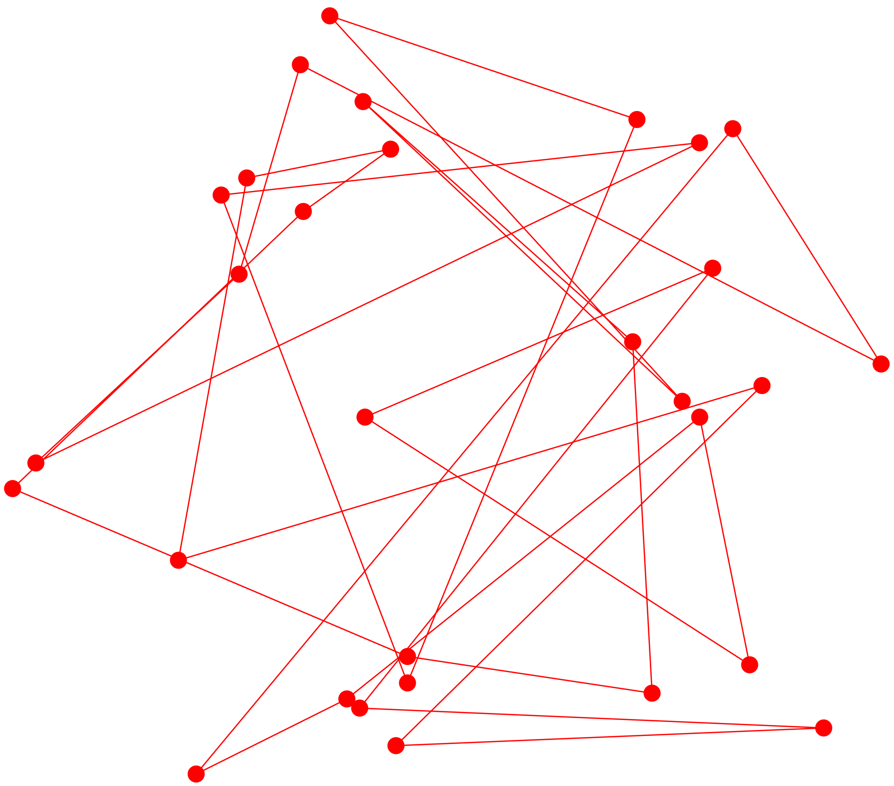
Repeat the process many times, show what happens graphically, and EXPLAIN WHAT YOU SEE.

The Iteration

```
% Random vectors with mean zero...

x = rand(n,1); x = x - mean(x); x = x/norm(x);
y = rand(n,1); y = y - mean(y); y = y/norm(y);

for i=1:nRepeat

   % Connect midpoints, scale, and plot...
   x = (x + [x(2:n);x(1)])/2;    x = x/norm(x);
   y = (y + [y(2:n);y(1)])/2;    y = y/norm(y);
   plot(x,y)

end
```

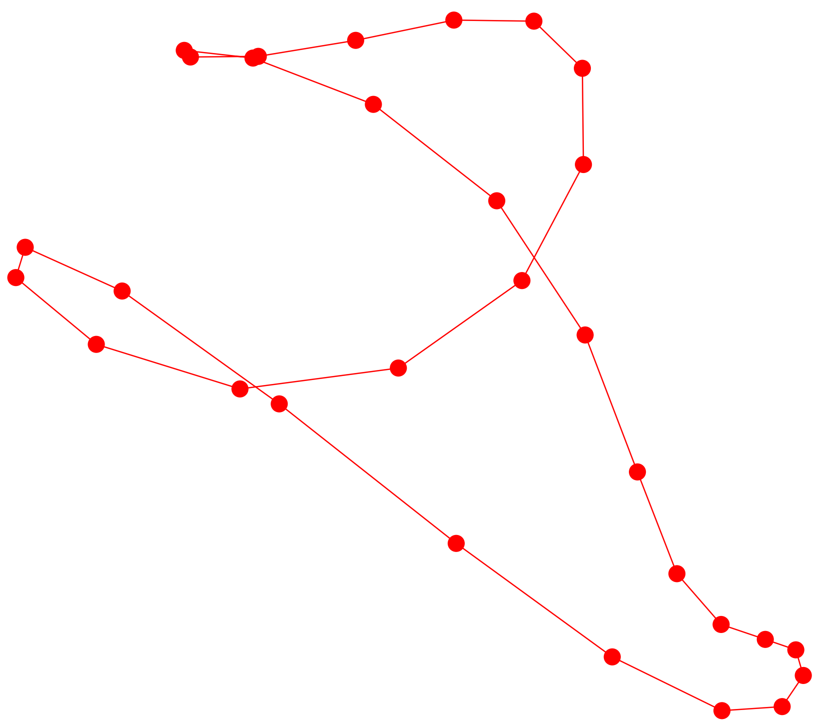
Solution 3

A Random Polygon (n = 30)



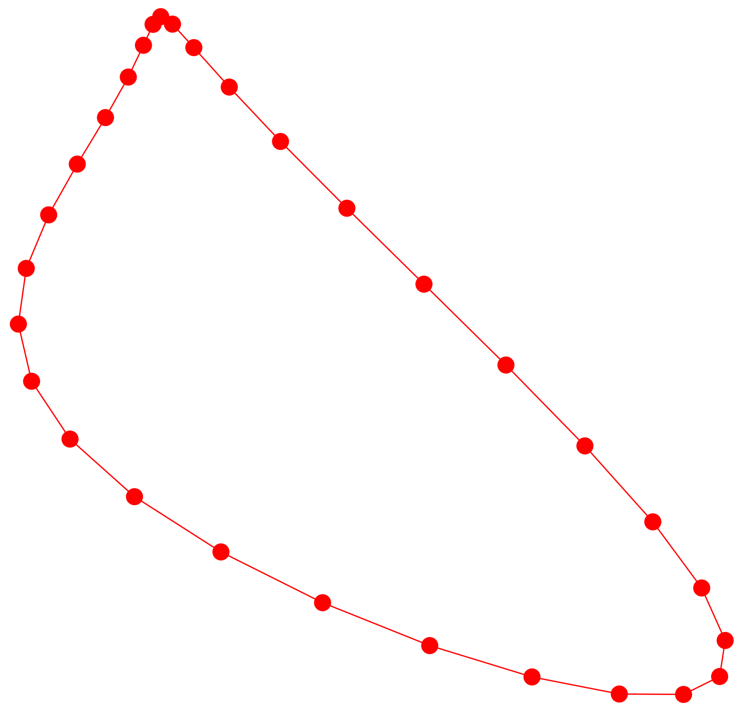
Solution 3

n = 30   Averagings = 20



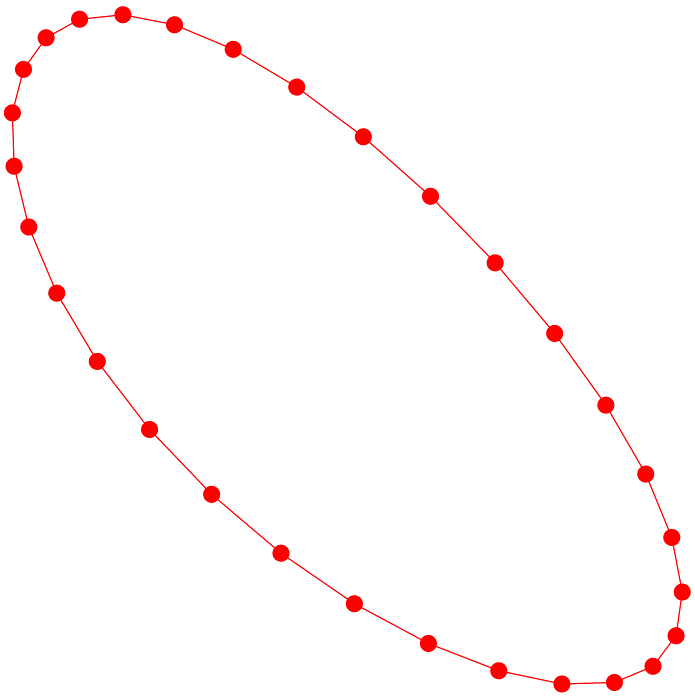
Solution 3

n = 30   Averagings = 100



Solution 3

n = 30   Averagings = 225



Matrix Description

Midpoint generation is a matrix-vector product computation...

$$\begin{bmatrix} \widehat{x}_1 \\ \widehat{x}_2 \\ \widehat{x}_3 \\ \widehat{x}_4 \\ \widehat{x}_5 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_3 + x_4 \\ x_4 + x_5 \\ x_5 + x_1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

$$\Uparrow$$

$$M_5$$

$$\Downarrow$$

$$\begin{bmatrix} \widehat{y}_1 \\ \widehat{y}_2 \\ \widehat{y}_3 \\ \widehat{y}_4 \\ \widehat{y}_5 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} y_1 + y_2 \\ y_2 + y_3 \\ y_3 + y_4 \\ y_4 + y_5 \\ y_5 + y_1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}$$

The Iteration

```
% Random vectors with mean zero...
x = rand(n,1); x = x - mean(x); x = x/norm(x);
y = rand(n,1); y = y - mean(y); y = y/norm(y);
for i=1:nRepeat
   % Connect midpoints, scale, and plot...
   x = M(n)*x;   x = x/norm(x);
   y = M(n)*y;   y = y/norm(y);
   plot(x,y)
end
```

Side-by-side instances of the power method.

Power Iteration

Suppose A is 5-by-5 and $Ax_i = \lambda_i x_i$ for $i = 1{:}5$. Assume

$$|\lambda_1| \; > \; |\lambda_2| \; > \; |\lambda_3| \; > \; |\lambda_4| \; > \; |\lambda_5|$$

If

$$v \; = \; a_1 \cdot x_1 \; + \; a_2 \cdot x_2 \; + \; a_3 \cdot x_3 \; + \; a_4 \cdot x_4 \; + \; a_5 \cdot x_5$$

then

$$Av \;= a_1\,\lambda_1 \cdot x_1 + a_2\,\lambda_2 \cdot x_2 + a_3\,\lambda_3 \cdot x_3 + a_4\,\lambda_4 \cdot x_4 + a_5\,\lambda_5 \cdot x_5$$

$$A^2v \;= a_1\,\lambda_1^2 \cdot x_1 + a_2\,\lambda_2^2 \cdot x_2 + a_3\,\lambda_3^2 \cdot x_3 + a_4\,\lambda_4^2 \cdot x_4 + a_5\,\lambda_5^2 \cdot x_5$$

$$A^kv \;= a_1\,\lambda_1^k \cdot x_1 + a_2\,\lambda_2^k \cdot x_2 + a_3\,\lambda_3^k \cdot x_3 + a_4\,\lambda_4^k \cdot x_4 + a_5\,\lambda_5^k \cdot x_5$$

Power Iteration

If

$$v = a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_3 + a_4 \cdot x_4 + a_5 \cdot x_5$$

then

$$A^k v = a_1 \lambda_1^k \cdot x_1 + a_2 \lambda_2^k \cdot x_2 + a_3 \lambda_3^k \cdot x_3 + a_4 \lambda_4^k \cdot x_4 + a_5 \lambda_5^k \cdot x_5$$

Focus on the *direction* of this vector:

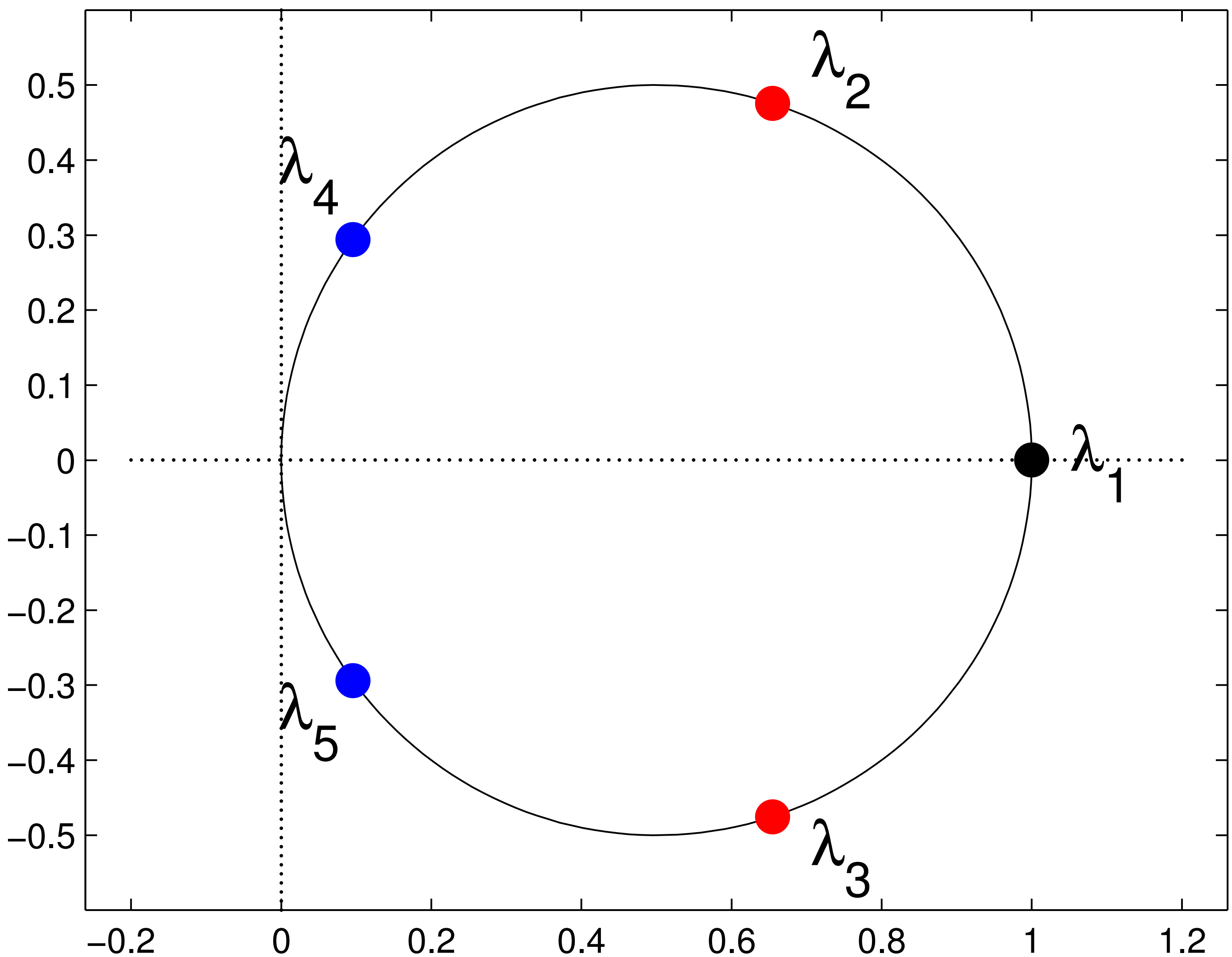
$$\lambda_1^k \left(a_1 x_1 + a_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k x_2 + a_3 \left(\frac{\lambda_3}{\lambda_1}\right)^k x_3 + a_4 \left(\frac{\lambda_4}{\lambda_1}\right)^k x_4 + a_5 \left(\frac{\lambda_5}{\lambda_1}\right)^k x_5 \right)$$

The Iteration

```
% Random vectors with mean zero...
x = rand(n,1); x = x - mean(x); x = x/norm(x);
y = rand(n,1); y = y - mean(y); y = y/norm(y);
for i=1:nRepeat
   % Connect midpoints, scale, and plot...
   x = M(n)*x;   x = x/norm(x);
   y = M(n)*y;   y = y/norm(y);
   plot(x,y)
end
```

Side-by-side instances of the power method.

The Eigenvalues of M_5



The Eigenspaces of M_5

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} \cos(0) \\ \cos(2\pi/5) \\ \cos(4\pi/5) \\ \cos(6\pi/5) \\ \cos(8\pi/5) \end{bmatrix}, \begin{bmatrix} \sin(0) \\ \sin(2\pi/5) \\ \sin(4\pi/5) \\ \sin(6\pi/5) \\ \sin(8\pi/5) \end{bmatrix} \right\}, \left\{ \begin{bmatrix} \cos(0) \\ \cos(4\pi/5) \\ \cos(8\pi/5) \\ \cos(12\pi/5) \\ \cos(16\pi/5) \end{bmatrix}, \begin{bmatrix} \sin(0) \\ \sin(4\pi/5) \\ \sin(8\pi/5) \\ \sin(12\pi/5) \\ \sin(16\pi/5) \end{bmatrix} \right\}$$

$$\lambda_1 = 1 \qquad |\lambda_2| = |\lambda_3| = \cos^2\left(\frac{\pi}{5}\right) \qquad |\lambda_4| = |\lambda_5| = \cos^2\left(\frac{2\pi}{5}\right)$$

A starting vector with mean 0 is orthogonal to the first eigenvector. $(a_1 = 0)$

Power Iteration

If

$$v \; = \; a_2 \cdot x_2 \; + \; a_3 \cdot x_3 \; + \; a_4 \cdot x_4 \; + \; a_5 \cdot x_5$$

then

$$A^k v \; = \; a_2 \, \lambda_2^k \cdot x_2 \; + \; a_3 \, \lambda_3^k \cdot x_3 \; + \; a_4 \, \lambda_4^k \cdot x_4 \; + \; a_5 \, \lambda_5^k \cdot x_5$$

Focus on the *direction* of this vector:

$$\lambda_2^k \left(a_2 \, x_2 \; + \; a_3 \left(\frac{\lambda_3}{\lambda_2} \right)^k x_2 \; + \; a_4 \left(\frac{\lambda_4}{\lambda_2} \right)^k x_3 \; + \; a_5 \left(\frac{\lambda_5}{\lambda_2} \right)^k x_4 \right)$$

The Damping Factor

$$\frac{|\lambda_4|}{|\lambda_2|} = \frac{\cos\left(\frac{2\pi}{n}\right)}{\cos\left(\frac{\pi}{n}\right)} = 1 - \frac{3}{2}\left(\frac{\pi}{n}\right)^2 + O\left(\frac{1}{n^4}\right)$$

After k averagings:

$$\text{dist(Vertices , Limiting Ellipse)} \approx \left(\frac{|\lambda_4|}{|\lambda_2|}\right)^k$$

At the Start...

$$x^{(0)} = \alpha_1 \begin{bmatrix} \cos\left(\frac{0\pi}{5}\right) \\ \cos\left(\frac{2\pi}{5}\right) \\ \cos\left(\frac{4\pi}{5}\right) \\ \cos\left(\frac{6\pi}{5}\right) \\ \cos\left(\frac{8\pi}{5}\right) \end{bmatrix} + \alpha_2 \begin{bmatrix} \sin\left(\frac{0\pi}{5}\right) \\ \sin\left(\frac{2\pi}{5}\right) \\ \sin\left(\frac{4\pi}{5}\right) \\ \sin\left(\frac{6\pi}{5}\right) \\ \sin\left(\frac{8\pi}{5}\right) \end{bmatrix} + \alpha_3 \begin{bmatrix} \cos\left(\frac{0\pi}{5}\right) \\ \cos\left(\frac{4\pi}{5}\right) \\ \cos\left(\frac{8\pi}{5}\right) \\ \cos\left(\frac{12\pi}{5}\right) \\ \cos\left(\frac{16\pi}{5}\right) \end{bmatrix} + \alpha_4 \begin{bmatrix} \sin\left(\frac{0\pi}{5}\right) \\ \sin\left(\frac{4\pi}{5}\right) \\ \sin\left(\frac{8\pi}{5}\right) \\ \sin\left(\frac{12\pi}{5}\right) \\ \sin\left(\frac{16\pi}{5}\right) \end{bmatrix}$$

$$y^{(0)} = \beta_1 \left[\quad \text{ditto} \quad \right] + \beta_2 \left[\quad \text{ditto} \quad \right] + \beta_3 \left[\quad \text{ditto} \quad \right] + \beta_4 \left[\quad \text{ditto} \quad \right]$$

In the Limit...

After k averagings, vertex vectors $x^{(k)}$ and $y^{(k)}$ are unit vectors in the direction of

$$M_5^k x^{(0)} \approx \alpha_1 M_5^k \begin{bmatrix} \cos\left(\frac{0\pi}{5}\right) \\ \cos\left(\frac{2\pi}{5}\right) \\ \cos\left(\frac{4\pi}{5}\right) \\ \cos\left(\frac{6\pi}{5}\right) \\ \cos\left(\frac{8\pi}{5}\right) \end{bmatrix} + \alpha_2 M_5^k \begin{bmatrix} \sin\left(\frac{0\pi}{5}\right) \\ \sin\left(\frac{2\pi}{5}\right) \\ \sin\left(\frac{4\pi}{5}\right) \\ \sin\left(\frac{6\pi}{5}\right) \\ \sin\left(\frac{8\pi}{5}\right) \end{bmatrix}$$

$$M_5^k y^{(0)} \approx \beta_1 M_5^k \left[\quad \text{ditto} \quad \right] + \beta_2 M_5^k \left[\quad \text{ditto} \quad \right]$$

What Those Unit Vectors Are.....

$$x^{(k)} \approx \cos(\theta_{\mathbf{x}}) \begin{bmatrix} \cos\left(\frac{(0+k)\pi}{5}\right) \\ \cos\left(\frac{(2+k)\pi}{5}\right) \\ \cos\left(\frac{(4+k)\pi}{5}\right) \\ \cos\left(\frac{(6+k)\pi}{5}\right) \\ \cos\left(\frac{(8+k)\pi}{5}\right) \end{bmatrix} + \sin(\theta_{\mathbf{x}}) \begin{bmatrix} \sin\left(\frac{(0+k)\pi}{5}\right) \\ \sin\left(\frac{(2+k)\pi}{5}\right) \\ \sin\left(\frac{(4+k)\pi}{5}\right) \\ \sin\left(\frac{(6+k)\pi}{5}\right) \\ \sin\left(\frac{(8+k)\pi}{5}\right) \end{bmatrix}$$

$$y^{(k)} \approx \cos(\theta_{\mathbf{y}}) \begin{bmatrix} \text{ditto} \end{bmatrix} + \sin(\theta_{\mathbf{y}}) \begin{bmatrix} \text{ditto} \end{bmatrix}$$

The Four Magic Numbers.....

$$\cos(\theta_x) \;=\; \frac{c^T x^{(0)}}{\sqrt{(c^T x^{(0)})^2 \;+\; (s^T x^{(0)})^2}}$$

$$c^T = \left[\; \cos\left(\frac{0\pi}{5}\right) \quad \cos\left(\frac{2\pi}{5}\right) \quad \cos\left(\frac{4\pi}{5}\right) \quad \cos\left(\frac{6\pi}{5}\right) \quad \cos\left(\frac{8\pi}{5}\right) \;\right]$$

$$s^T = \left[\; \sin\left(\frac{0\pi}{5}\right) \quad \sin\left(\frac{2\pi}{5}\right) \quad \sin\left(\frac{4\pi}{5}\right) \quad \sin\left(\frac{6\pi}{5}\right) \quad \sin\left(\frac{8\pi}{5}\right) \;\right]$$

The recipes for $\sin(\theta_x)$, $\cos(\theta_y)$, and $\sin(\theta_y)$ are similar.

Where the Vertices Sit.....

$$x(t) \approx \cos(\theta_{\mathbf{x}})\cos(t) + \sin(\theta_{\mathbf{x}})\sin(t)$$

$$y(t) \approx \cos(\theta_{\mathbf{y}})\cos(t) + \sin(\theta_{\mathbf{y}})\sin(t)$$

This describes an ellipse!

What Are Its Semiaxes and Tilt?

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \approx \begin{bmatrix} \cos(\theta_{\mathbf{x}}) & \sin(\theta_{\mathbf{x}}) \\ \cos(\theta_{\mathbf{y}}) & \sin(\theta_{\mathbf{y}}) \end{bmatrix} \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$$

$$= U \begin{bmatrix} \sigma_{\mathbf{1}} & \mathbf{0} \\ \mathbf{0} & \sigma_{\mathbf{2}} \end{bmatrix} V^T \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$$

The SVD Tells All...

2-by-2 Singular Value Decomposition

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \cos(\theta_u) & \sin(\theta_u) \\ -\sin(\theta_u) & \cos(\theta_u) \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} \cos(\theta_v) & \sin(\theta_v) \\ -\sin(\theta_v) & \cos(\theta_v) \end{bmatrix}^T$$

The ellipse connection:

$$\mathcal{E} = \left\{ (x(t), y(t)) \;\middle|\; \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} \right\}$$

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \cos(t-\theta_v)\sigma_1 \begin{bmatrix} \cos(\theta_u \\ -\sin(\theta_v \end{bmatrix} + \sin(t-\theta_v)\sigma_1 \begin{bmatrix} \sin(\theta_u \\ \cos(\theta_v \end{bmatrix}$$

What Are Its Semiaxes and Tilt?

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \approx \begin{bmatrix} \cos(\theta_{\mathbf{x}}) & \sin(\theta_{\mathbf{x}}) \\ \cos(\theta_{\mathbf{y}}) & \sin(\theta_{\mathbf{y}}) \end{bmatrix} \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$$

$$= U \begin{bmatrix} \sigma_{\mathbf{1}} & \mathbf{0} \\ \mathbf{0} & \sigma_{\mathbf{2}} \end{bmatrix} V^T \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$$

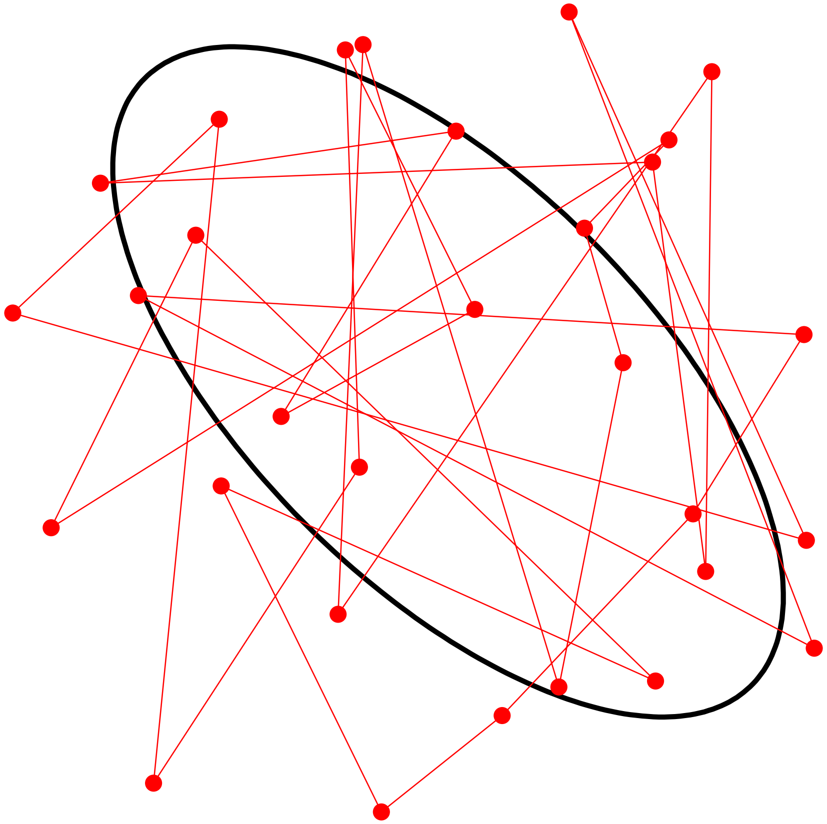
The SVD Tells All...

The Limiting Ellipse

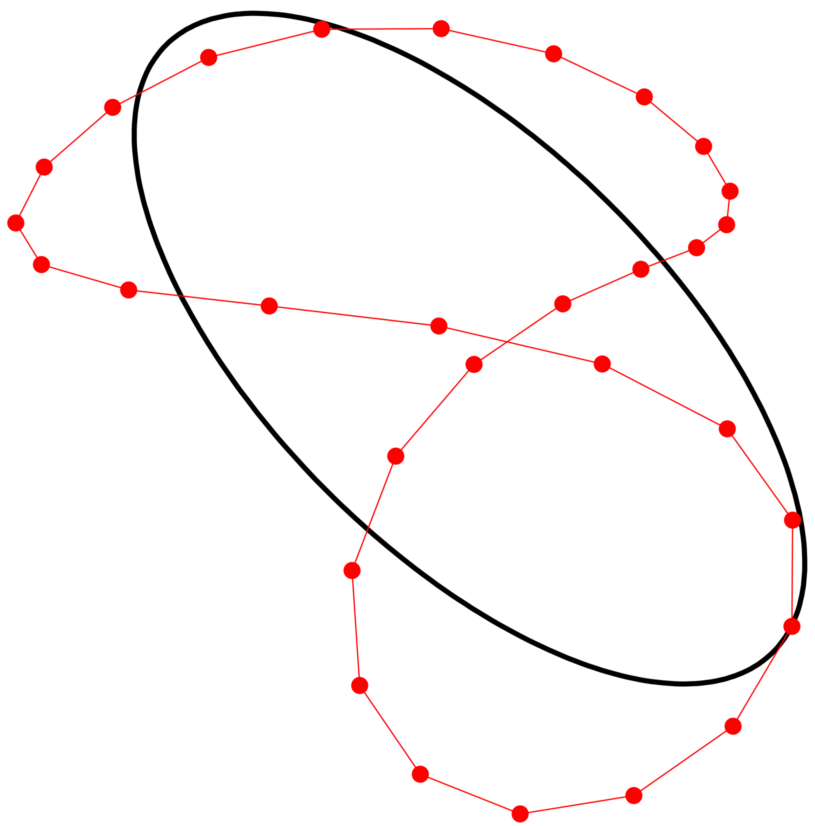
$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \approx \begin{bmatrix} \cos(\mathbf{45}) & -\sin(\mathbf{45}) \\ \cos(\mathbf{45}) & \sin(\mathbf{45}) \end{bmatrix} \begin{bmatrix} \sigma_{\mathbf{1}} & \mathbf{0} \\ \mathbf{0} & \sigma_{\mathbf{2}} \end{bmatrix} \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$$

$$\sigma_1 = \frac{2}{\sqrt{n}} \cdot \cos\left(\frac{\theta_y - \theta_x}{2}\right) \qquad \sigma_2 = \frac{2}{\sqrt{n}} \cdot \sin\left(\frac{\theta_y - \theta_x}{2}\right)$$

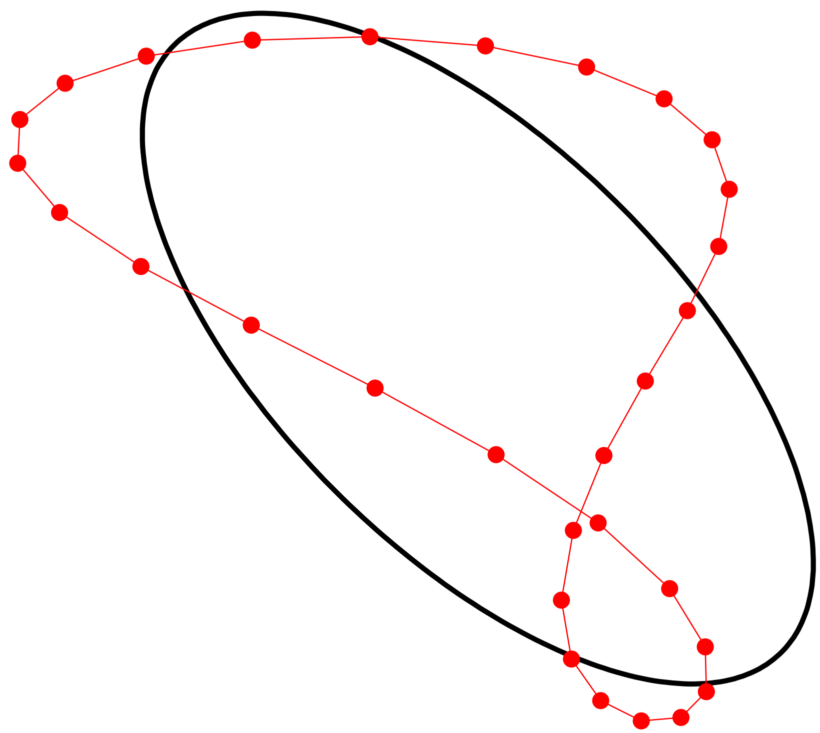
An Initial "Random" Polygon



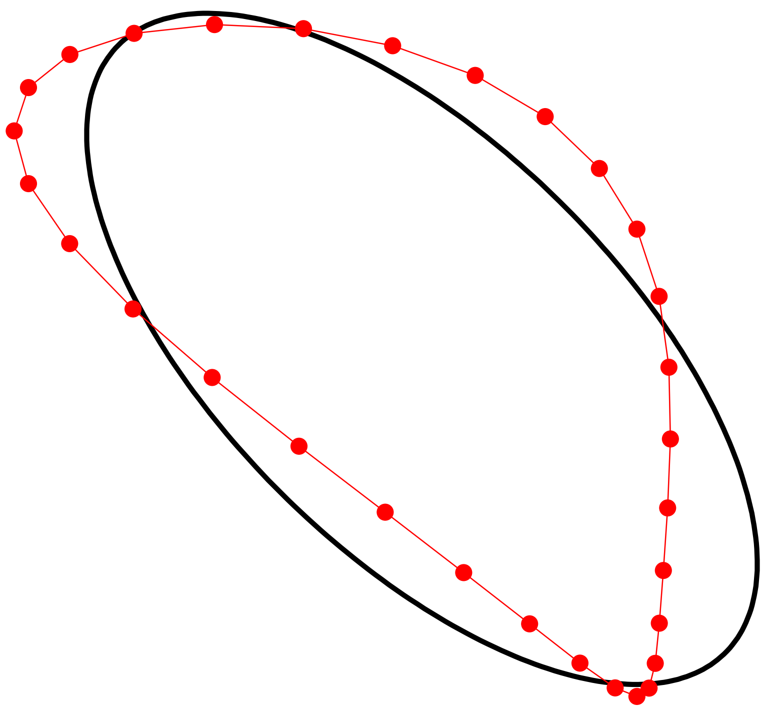
After 32 Averagings...



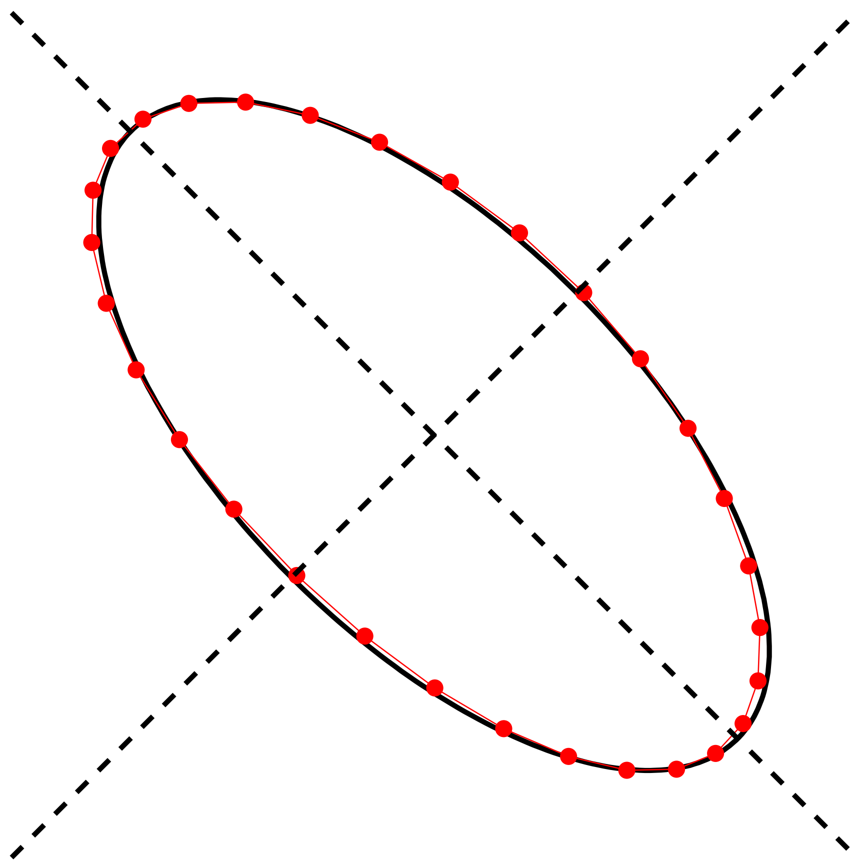
After 80 Averagings...



After 140 Averagings...

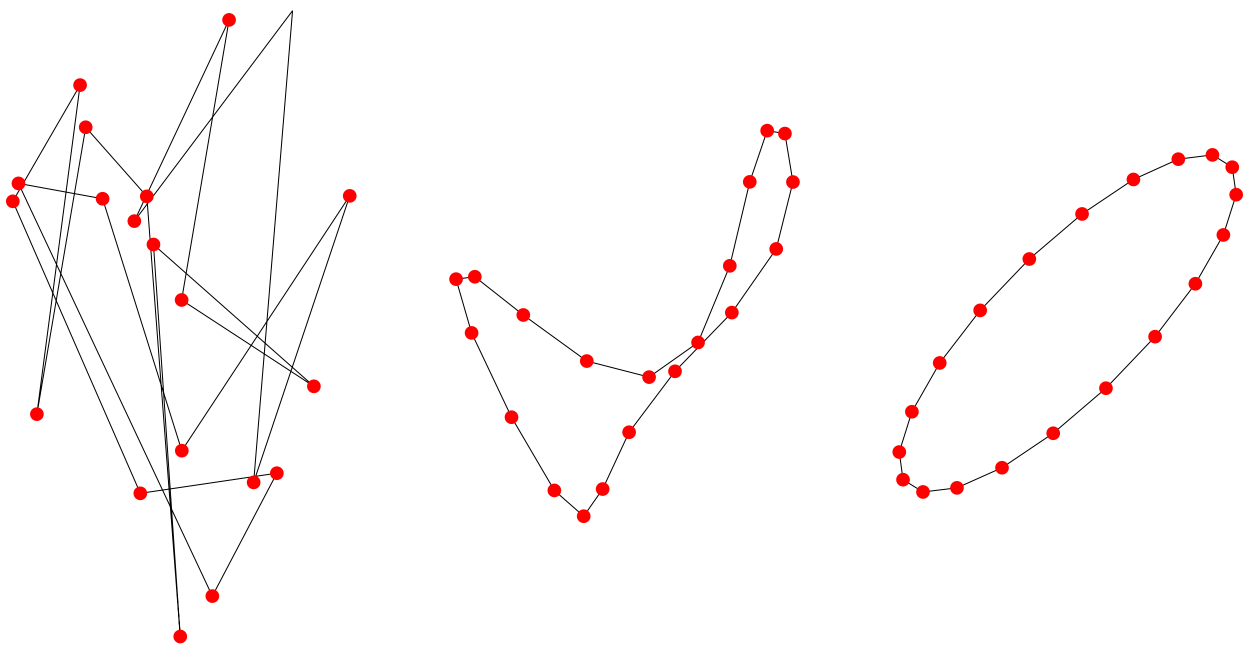


After 278 Averagings...



An Undergraduate Research Problem

Explain why the limiting ellipse has a 45-degree tilt? What are the lengths of its semiaxes?

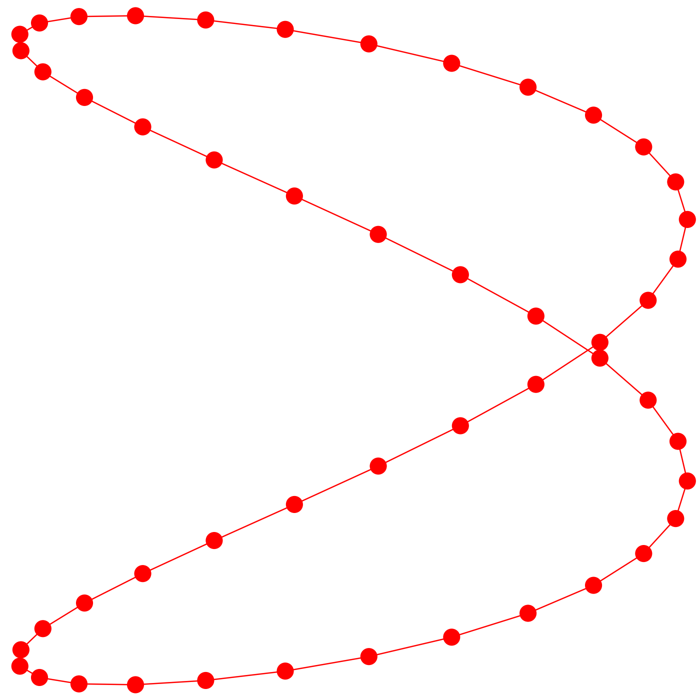


A. Elmachtaub and C. Van Loan (2010). "From Random Polygon to Ellipse," *SIAM Review*, 52, 151-170.

When the Initial Vertex Vectors Are Deficient...

x orthoginal to D2

n = 50 Averagings = 2328



More Interesting Behavior...

1. What happens if we scale by other norms? E.g.,

$$x = x / \| x \|_p$$

2. What happens if we use alternative "midpoints"? E.g.,

$$x_i = \lambda \cdot x_i \ + \ (1 - \lambda) \cdot x_{i+1}$$

The Problem is a Metaphor for Computational Science and Engineering

First:

We *experimented* with a simple iteration and observed that it transforms something that is chaotic and rough into something that is organized and smooth.

The Problem is a Metaphor for Computational Science and Engineering

Second:

As a step towards explaining the limiting behavior of the polygon sequence, we described the averaging process using *matrix-vector notation*.

The Problem is a Metaphor for Computational Science and Engineering

Third:

This led to an *eigenanalysis*, the identification of a crucial *invariant subspace*, and a vertex-vector *convergence analysis*.

The Problem is a Metaphor for Computational Science and Engineering

Fourth:

We then used the *singular value decomposition* to connect our algebraic manipulations to a simple underlying geometry.

Conclusions

1. It is never too early for big ideas to show up in the ugrad curriculum.

2. An important big idea is the use of matrix computations to explain phenomena.