# Normalized Iterated Averaging Polygons

Yonah Biers-Ariel[*], Matthew Charnley[*†], Brooke Logan[*], Anthony Zaleski[*], Doron Zeilberger[‡]

June 13, 2018

**Abstract**

Inspired by the beautiful article by Adam Elmachtoub and Charles Van Loan, we take an initial polygon consisting of $n$ points in the Cartesian plane with centroid at the origin and define an averaging procedure, generating a new polygon out of the midpoints of each segment that defines the initial polygon. We perform an analysis of this procedure, corroborating previous results on this procedure when the polygon is normalized at each step. We also introduce a new averaging procedure suggested by the power method which also converges to an ellipse but can be in any orientation. We conclude by considering other averaging procedures and analyze the results that come from these generalizations as well.

## 1 Introduction

We consider a polygon defined by $n$ points in the plane, where the points are connected in order, possibly resulting in a non-convex shape with self-intersections. As discussed in [3], which refers to the paper [2], iterating an averaging procedure on these polygons causes the points to all lie on an ellipse. Numerical calculations performed using Maple also verify this fact; the points after each step of the iteration converge to all lie on an ellipse, and the points all converge to the origin. In order to combat this, the authors in [2] introduce a normalizing procedure to keep the ellipse at a non-zero size so that they can analyze the resulting ellipse. In particular, they find that for any initial set of points, the result under this procedure is an ellipse that is tilted at a 45 degree angle from the coordinate axes. However, our initial averaging procedures resulted in ellipses that were not only tilted at a 45 degree angle, so we set out to perform another analysis of this procedure.

In this paper, we reconstruct the averaging procedure defined in [2] and analyze the results. We verify the results presented in that paper following their normalization scheme. After that, we present a new normalization scheme motivated by the power method of numerical analysis that also serves to retain the size of the the polygon as the iteration continues, giving a different set of ellipses as possible limiting solutions. We conclude this paper with further generalizations of the averaging procedure and discover what kind of limiting shapes can be achieved. In addition to the proofs of these results, Maple code is provided that iterates these procedures and shows the results.

## 2 Averaging Polygons

Let $P$ be a polygon in the plane $\mathbb{R}^2$, denoted by its set of vertices

$$\{(x_0, y_0), ..., (x_{n-1}, y_{n-1})\},$$

[*]Members of Dr. Zeilberger's Spring 2018 Experimental Mathematics Class
[†]Corresponding author: charnley@math.rutgers.edu
[‡]Rutgers University, doronzeil@gmail.com

where the vertices are connected in order, with $(x_{n-1}, y_{n-1})$ connected to $(x_0, y_0)$. In this definition, the polygon does not have to be convex, and the edges can self-intersect. By translating the polygon, we will assume that the centroid is the origin, namely that

$$\sum_{i=0}^{n-1} x_i = 0 = \sum_{i=0}^{n-1} y_i.$$

As discussed in [2], we will consider an averaging process on the vertices of the polygon, which will consist of constructing a new polygon $\tilde{P}$ from $P$ by defining

$$\tilde{x}_i = \frac{1}{2}(x_i + x_{i+1}), \qquad \tilde{y}_i = \frac{1}{2}(y_i + y_{i+1}), \qquad i = 0, ..., n-2$$

and

$$\tilde{x}_{n-1} = \frac{1}{2}(x_{n-1} + x_0), \qquad \tilde{y}_{n-1} = \frac{1}{2}(y_{n-1} + y_0).$$

Therefore, if we define the vector $\mathbf{x}$ to be the list of $x$-components of the points and $\mathbf{y}$ to be the corresponding $y$-components, we can see that this iteration is defined by

$$\tilde{\mathbf{x}} = A\mathbf{x} \qquad \tilde{\mathbf{y}} = A\mathbf{y}$$

for the matrix

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2} & 0 & 0 & 0 & \cdots & \frac{1}{2} \end{bmatrix},$$

so that iterating this averaging procedure is equivalent to iterating the matrix $A$ on both the $x$ and $y$ components of the polygon. Thus, to determine the long-term behavior of this procedure, we should analyze the eigenvalues and eigenvectors of this matrix $A$.

**Lemma 2.1.** *The eigenvalues of $A$ are*

$$\lambda_k := \frac{1}{2} + \frac{\omega^k}{2}, \; k = 0, 1, ..., n-1 \qquad \text{where } \omega = e^{\frac{2\pi i}{n}},$$

*and the corresponding eigenvector is $\mathbf{v}_k := \frac{1}{\sqrt{n}}[1, \omega^k, \omega^{2k}, ..., \omega^{(n-1)k}]^T$ for each $k = 0, 1, ..., n-1$. This is an orthonormal set of eigenvectors for this matrix.*

*Proof.* We can write the matrix $A$ as

$$A = \frac{1}{2}I + \frac{1}{2}S,$$

where $S$ is the shift matrix given by

$$S = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

By induction, we can prove that the characteristic polynomial of $S$ is $\lambda^n - 1$, so that the eigenvalues are the $n$th roots of unity, $\{\omega^k\}$. Futhermore, a simple computation shows the normalized eigenvectors of $S$ are exactly of the form $\mathbf{v}_k$ from the statement. Since the identity matrix preserves every vector, we see that these $\mathbf{v}_k$ are still eigenvectors for $A$, and their corresponding eigenvalues are given by the formula $\lambda_k$ above. We can see that these vectors are orthonormal by noticing that

$$\overline{\mathbf{v}_0} = \mathbf{v}_0, \qquad \overline{\mathbf{v}_k} = \mathbf{v}_{n-k},$$

and so

$$\mathbf{v}_j \cdot \overline{\mathbf{v}_j} = \mathbf{v}_j \cdot \mathbf{v}_{n-j} = \frac{1}{n}(1 + \omega^n + \omega^{2n} + \cdots + \omega^{(n-1)n}) = 1,$$

so that each vector has norm 1. They are also orthogonal because the elements in the sum of $\mathbf{v}_j \cdot \overline{\mathbf{v}_k}$ are the same as the elements of the eigenvector corresponding to $\mathbf{v}_{j+n-k \mod n}$, which is zero by the next lemma if $j \neq k$. ☐

**Lemma 2.2.** *Let $\mathbf{v}_k$ be an eigenvector of $A$ for $1 \leq k \leq n-1$. Then the sum of the components of $\mathbf{v}_k$ is zero.*

*Proof.* If $k \neq 0$, the elements of $\mathbf{v}_k$ are all roots of unity, and the components of each vector is a geometric progression starting with 1 and $\omega^k$. Therefore, we can consider the group $\mathbb{Z}_n$, integers modulo $n$ under modular addition, and the subgroup $G_k$ generated by the element $k$, and the exponents of $\omega$ in the vector $\mathbf{v}_k$ will be exactly the elements in $G_k$, appearing in the order given by $\{0, k, 2k, ...\}$. If $|G_k| = n$, then the elements in $\mathbf{v}_k$ are exactly the $n$th roots of unity, and so their sum is zero. If $|G_k| < n$, then $|G_k| = l$, which divides $n$, and so the elements of $\mathbf{v}_k$ will correspond to the $l$th roots of unity, each appearing $\frac{n}{l}$ times. Therefore, the sum of the elements of $\mathbf{v}_k$ is the same as $\frac{n}{l}$ times the sum of the $l$th roots of unity, which is still zero. Therefore, the sum in each case is zero. ☐

**Corollary 2.1.** *Let $w$ be any vector in $\mathbb{R}^n$. If $w$ is decomposed according to the eigenvectors of $A$ in the form*

$$w = \sum_{k=0}^{n-1} \alpha_k \mathbf{v}_k,$$

*then $\alpha_0 = 0$ if and only if the sum of the components of $w$ is zero.*

*Proof.* Since the sum of the components of $\mathbf{v}_k = 0$ for all $k > 0$, the sum of the components of $w$ is equal to $\alpha_0$ times $\sqrt{n}$, which is the sum of components of $v_0$. ☐

**Theorem 2.1.** *If $P$ is a polygon with centroid at the origin, then the vectors $x$ and $y$ have a decomposition of the form*

$$\mathbf{x} = \sum_{k=1}^{n-1} \alpha_k \mathbf{v}_k \qquad \mathbf{y} = \sum_{k=1}^{n-1} \beta_k \mathbf{v}_k, \tag{1}$$

*where $\alpha_k = \mathbf{x}^T \overline{\mathbf{v}_k}$ and $\beta_k = \mathbf{y}^T \overline{\mathbf{v}_k}$.*

*Proof.* Since the components of the vectors $\mathbf{x}$ and $\mathbf{y}$ add to zero, the previous corollary implies that they have no coefficient on the vector $v_0$. Furthermore, since the $\{\mathbf{v}_k\}$ are orthonormal, we can find the coefficients by taking the dot product with the $\mathbf{x}$ or $\mathbf{y}$ vectors. ☐

If we look at iterating the matrix $A$, the power method of numerical analysis [1] tells us that the dominant component in the limit of $A^m w$ is the part with the largest eigenvalue in magnitude. In this case, the eigenvalue of $A$ with the largest eigenvalue in magnitude has $|\lambda_0| = 1$, but that eigenvector is not present in our polygons because they have centroid at the origin. Therefore, we have the following

**Lemma 2.3.** *For each of the vectors $\mathbf{x}$ and $\mathbf{y}$ from a polygon $P$ with eigenvector decompositions given in* (1)*, we have that as $m \to \infty$,*

$$A^m \mathbf{x} = |\lambda_1|^m \left( \left( \frac{\lambda_1}{|\lambda_1|} \right)^m \alpha_1 \mathbf{v}_1 + \left( \frac{\lambda_{n-1}}{|\lambda_{n-1}|} \right)^m \alpha_{n-1} \mathbf{v}_{n-1} + O\left( \left( \frac{|\lambda_2|}{|\lambda_1|} \right)^m \right) \right),$$

$$A^m\mathbf{y} = |\lambda_1|^m\left(\left(\frac{\lambda_1}{|\lambda_1|}\right)^m\beta_1\mathbf{v}_1 + \left(\frac{\lambda_{n-1}}{|\lambda_{n-1}|}\right)^m\beta_{n-1}\mathbf{v}_{n-1} + O\left(\left(\frac{|\lambda_2|}{|\lambda_1|}\right)^m\right)\right).$$

*Proof.* If $\mathbf{x}$ is of the form in (1), then the vector $A^m\mathbf{x}$ is given by

$$\sum_{i=1}^{n-1}\lambda_i^m\alpha_i\mathbf{v}_i,$$

and similarly for the $y$ vector. By manipulating this formula, factoring out $|\lambda_1|^m$, we get the formulas above. In order to get something useful out of the formula, we need to establish that $|\lambda_1| = |\lambda_{n-1}|$ and this is the largest remaining eigenvalue. Since the eigenvalues of $A$ are of the form

$$\frac{1}{2} + \frac{\omega^k}{2}$$

the points in the complex plane corresponding to the $\lambda_k$ lie on a circle of radius $\frac{1}{2}$ centered at $\frac{1}{2}$. In order to find the eigenvalues of largest magnitude, we need to look for the ones farthest from the origin, which would be the ones in which the $\omega^k$ part of the expression is closest to being real-valued. This corresponds to $k = 1$ and $k = n-1$ because $k = 0$ is on the real axis, but is not present in our polygons. Furthermore, since $\lambda_1$ and $\lambda_{n-1}$ are complex conjugates, they have the same magnitude. Since these have the largest eigenvalue, the big-O term will go to zero as $m \to \infty$, and so the above formulas hold and are useful for doing a limit analysis on this process. □

Therefore, if we want to look at the long-term behavior of this iteration process, we need to consider the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_{n-1}$. Since these correspond to roots of unity, we know that $\mathbf{v}_{n-1} = \overline{\mathbf{v}_1}$. This implies that if $\mathbf{x}$ and $\mathbf{y}$ are real vectors with eigenvector decomposition given by (1), then we must have $\alpha_{n-1} = \overline{\alpha_1}$, so that

$$\alpha_1\mathbf{v}_1 + \alpha_{n-1}\mathbf{v}_{n-1} = \alpha_1\mathbf{v}_1 + \overline{\alpha_1\mathbf{v}_1} = 2\Re(\alpha_1\mathbf{v}_1),$$

and similarly for $\beta_1$ and $\beta_{n-1}$. Since $\lambda_{n-1} = \overline{\lambda_1}$, we can modify Lemma 2.3 to

**Lemma 2.4.** *For each of the vectors* $\mathbf{x}$ *and* $\mathbf{y}$ *from a polygon* $P$ *with eigenvector decompositions given in* (1)*, we have that*

$$\lim_{m\to\infty} A^m\mathbf{x} = |\lambda_1|^m\left(2\Re\left(\frac{\lambda_1^m}{|\lambda_1|^m}\alpha_1\mathbf{v}_1\right) + O\left(\left(\frac{|\lambda_2|}{|\lambda_1|}\right)^m\right)\right),$$

$$\lim_{m\to\infty} A^m\mathbf{y} = |\lambda_1|^m\left(2\Re\left(\frac{\lambda_1^m}{|\lambda_1|^m}\beta_1\mathbf{v}_1\right) + O\left(\left(\frac{|\lambda_2|}{|\lambda_1|}\right)^m\right)\right).$$

Therefore, the shape corresponding to $\Re(\alpha_1\mathbf{v}_1)$ and $\Re(\beta_1\mathbf{v}_1)$ should have something to do with the limiting shape of this procedure. We will first analyze this shape, and then prove that this is sufficient for determining the limiting shape of the entire iterating process.

Consider a complex number $z = \sigma + i\tau$ and the vector $\mathbf{v}_1$ as before. Then we have that

$$\begin{aligned}
\Re(z\mathbf{v}_1) &= \Re\left((\sigma + i\tau)\frac{1}{\sqrt{n}}\begin{bmatrix} 1 \\ \omega \\ \vdots \\ \omega^{n-1}\end{bmatrix}\right) \\
&= \frac{1}{\sqrt{n}}\Re\left((\sigma+i\tau)\begin{bmatrix} \cos(0) + i\sin(0) \\ \cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right) \\ \vdots \\ \cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right)\end{bmatrix}\right) \\
&= \frac{1}{\sqrt{n}}\sigma\begin{bmatrix} \cos(0) \\ \cos\left(\frac{2\pi}{n}\right) \\ \vdots \\ \cos\left(\frac{2\pi}{n}\right)\end{bmatrix} - \frac{1}{\sqrt{n}}\tau\begin{bmatrix} \sin(0) \\ \sin\left(\frac{2\pi}{n}\right) \\ \vdots \\ \sin\left(\frac{2\pi}{n}\right)\end{bmatrix}.
\end{aligned}$$

Therefore, all of the points of the vector $\Re(z\mathbf{v}_1)$ are of the form $\frac{\sigma}{\sqrt{n}}\cos(t_j) - \frac{\tau}{\sqrt{n}}\sin(t_j)$ for prescribed values of $t_j$. Applying this to vectors for $\mathbf{x}$ and $\mathbf{y}$, we have the following.

**Theorem 2.2.** *Let $\hat{\mathbf{x}} = \Re(\alpha_1\mathbf{v}_1)$ and $\hat{\mathbf{y}} = \Re(\beta_1\mathbf{v}_1)$. Then all points of the form $(\hat{\mathbf{x}}_j, \hat{\mathbf{y}}_j)$ lie on the curve parametrized by*

$$\frac{1}{\sqrt{n}}(2\Re(\alpha_1)\cos(t) - 2\Im(\alpha_1)\sin(t), 2\Re(\beta_1)\cos(t) - 2\Im(\beta_1)\sin(t)), \qquad 0 \le t \le 2\pi.$$

**Lemma 2.5.** *If $\Re(z\mathbf{v}_1)$ is a set of points of the form $A\cos(t_j) + B\sin(t_j)$ for real numbers $A$ and $B$, and prescribed values of $t_j$, then the set of points given by $\Re(e^{i\phi}z\mathbf{v}_1)$ is of the same form, with the same $A$ and $B$, but potentially different values of $t_j$*

*Proof.* If we group the $e^{i\phi}$ into the $\mathbf{v}_1$ term, we get that

$$e^{i\phi}\mathbf{v}_1 = e^{i\phi}\begin{bmatrix} 1 \\ \omega \\ \vdots \\ \omega^{n-1}\end{bmatrix} = \begin{bmatrix} e^{i\phi} \\ e^{i(\frac{2\pi}{n}+\phi)} \\ \vdots \\ e^{i(\frac{2\pi(n-1)}{n}+\phi)}\end{bmatrix}.$$

This vector can be split into sine and cosine terms like before, and all that has changed is that the input to the sine and cosine functions has increase by $\phi$. Carrying the analysis through gives that all of these points lie on a curve of the same form, but the $t$ values have been shifted up by $\phi$. ☐

**Corollary 2.2.** *The points defined by $\Re(z\mathbf{v}_1)$ and $\Re(\left(\frac{\lambda_1}{|\lambda_1|}\right)^m z\mathbf{v}_1)$ all lie on the curve $\frac{\sigma}{\sqrt{n}}\cos(t) - \frac{\tau}{\sqrt{n}}\sin(t)$. Therefore the curve defined in the above contains all points in the set of the iterates $\Re(\left(\frac{\lambda_1}{|\lambda_1|}\right)^m z\mathbf{v}_1)$, in addition to the points in the vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$.*

## 2.1 Ellipse Analysis

**Lemma 2.6.** *Consider the curve $\gamma(t)$ given by*

$$(A\cos(t) + B\sin(t), C\cos(t) + D\sin(t)), \qquad 0 \le t \le 2\pi.$$

*Then, assuming $AD - BC \neq 0$ the curve $\gamma(t)$ traces out the ellipse given in normal form as*

$$Gx^2 + Hxy + Ky^2 = 1,$$

*with*

$$G = \frac{D^2 + C^2}{(AD - BC)^2}, \qquad H = \frac{2BD + 2AC}{(AD - BC)^2}, \qquad K = \frac{A^2 + B^2}{(AD - BC)^2}.$$

*Proof.* Setting $x = A\cos(t) + B\sin(t)$ and $y = C\cos(t) + D\sin(t)$ in the desired form

$$Gx^2 + Hxy + Ky^2 = 1$$

gives an equation involving $\cos^2(t)$, $\cos(t)\sin(t)$ and $\sin^2(t)$. In order to get this to add to 1, we want the coefficients of $\cos^2(t)$ and $\sin^2(t)$ to be 1, and the coefficient of $\cos(t)\sin(t)$ to be zero. This gives rise to a linear system of equation of the form

$$\begin{bmatrix} A^2 & AC & C^2 \\ 2AB & AD + BC & 2CD \\ B^2 & BD & D^2 \end{bmatrix} \begin{bmatrix} G \\ H \\ K \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Solving this equation gives the values of $G$, $H$ and $K$ given above. □

*Remark.* If $AD - BC = 0$, then we get that the $y$ equation above is a constant multiple of the $x$ equation, so that the points traced out by $\gamma$ all lie on a line segment. This will generally not happen if the points are in general position, so we ignore this case.

**Lemma 2.7.** *If the ellipse given by*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

*is rotated by an angle of $\theta$, then the resulting ellipse is given in normal form by the expression*

$$\left(\frac{\cos^2(\theta)}{a^2} + \frac{\sin^2(\theta)}{b^2}\right)x^2 + 2\cos(\theta)\sin(\theta)\left(\frac{1}{a^2} - \frac{1}{b^2}\right)xy + \left(\frac{\sin^2(\theta)}{a^2} + \frac{\cos^2(\theta)}{b^2}\right)y^2 = 1.$$

*Proof.* Sending $x \mapsto x\cos(\theta) - y\sin(\theta)$ and $y \mapsto y\cos(\theta) + x\sin(\theta)$ in the standard equation of the ellipse and simplifying the expression gives the desired formula. □

*Remark.* (a) The values for $\theta$, $a$, and $b$ in the above expression can be found by first noting that

$$\tan(2\theta) = \frac{H}{(G - K)},$$

and then using this value of $\theta$ to compute

$$\frac{1}{a^2} = \frac{1}{2}\left(G + K + \frac{H}{\sin(2\theta)}\right), \qquad \frac{1}{b^2} = \frac{1}{2}\left(G + K - \frac{H}{\sin(2\theta)}\right).$$

(b) If we assume that $a \geq b$, then $\theta$ can be any value between 0 and $\pi$. However, if we allow either $a$ or $b$ to be the semi-major axis, then we can restrict $\theta$ to lie between 0 and $\pi/2$, and the sign of $H$ will tell us which is the semi-major axis. It turns out to be easier, from a numerical setting, to force $a > b$ and work from there. This mainly comes from the fact that when computing the inverse tangent, the result is always between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$.

(c) This $\theta$ is the angle needed to rotate the ellipse back to the standard orientation on the axes. Therefore, to get the angle that the ellipse was rotated, we need the opposite of the $\theta$ from the formula.

(d) The eccentricity of this ellipse can then calculated as

$$e = \sqrt{1 - \frac{b}{a}}.$$

**Corollary 2.3.** *If the curve $\gamma(t)$ is given by*

$$\gamma(t) = (A\cos(t) + B\sin(t), C\cos(t) + D\sin(t)), \qquad 0 \le t \le 2\pi,$$

*then $\gamma(t)$ traces out an ellipse tilted at a 45 degree angle if and only if $A^2 + B^2 = C^2 + D^2$.*

*Proof.* This happens if and only if $2\theta = \pi/2$, i.e., if $\tan(2\theta)$ is undefined, which happens when $G = K$, as seen in the formulas above. ☐

**Corollary 2.4.** *Consider the iterated polygon averaging procedure. The limiting curve is an ellipse in all cases (assuming the points are in general position). This ellipse is tilted at a 45 degree angle if and only if $|\alpha_1| = |\beta_1|$, where $\alpha_1$ and $\beta_1$ are defined by the eigenvector expansions* (1) *for the initial vectors* $\mathbf{x}$ *and* $\mathbf{y}$.

*Proof.* In our setup for the iterating procedure without normalization (normalzation will be discussed in the next section), $|\alpha_1|^2 = \frac{4}{n}(A^2 + B^2)$ and $|\beta_1|^2 = \frac{4}{n}(C^2 + D^2)$. Therefore, $A^2 + B^2 = C^2 + D^2$ if and only if $|\alpha_1| = |\beta_1|$. The result then follows from the previous corollary. ☐

# 3 Renormalization Considerations

Since $|\lambda_2| < 1$, we see that the averaging process applied to any polygon will converge to the origin, since both vectors $x$ and $y$ will go to zero. In order to visualize and control this process, we want to renormalize the vectors to get something that stays bounded. We will consider two methods of doing this, one described in [2] and another of our own design.

## 3.1 2-norm Normalization

In [2], the authors use an iteration process similar to ours. However, after each step of multiplying by the matrix $A$, they normalize both the $x$ and $y$ vectors to be of unit length. If $x$ and $y$ can be written in the form (1), then this process becomes

$$\tilde{x} = \frac{Ax}{\|Ax\|_2} = \frac{\sum_{i=1}^{n-1} \lambda_i \alpha_i \mathbf{v}_i}{\left(\sum_{j=1}^{n-1} |\lambda_j \alpha_j|^2\right)^{1/2}}, \qquad \tilde{y} = \frac{Ay}{\|Ay\|_2} = \frac{\sum_{i=1}^{n-1} \lambda_i \beta_i \mathbf{v}_i}{\left(\sum_{j=1}^{n-1} |\lambda_j \beta_j|^2\right)^{1/2}}.$$

If we, with motivation from the power method, factor a $|\lambda_1|$ from the numerator and denominator, we get

$$\tilde{x} = \frac{\sum_{i=1}^{n-1} \frac{\lambda_i}{|\lambda_1|} \alpha_i \mathbf{v}_i}{\left(\sum_{j=1}^{n-1} \frac{1}{|\lambda_1|^2} |\lambda_j \alpha_j|^2\right)^{1/2}}, \qquad \tilde{y} = \frac{\sum_{i=1}^{n-1} \frac{\lambda_i}{|\lambda_1|} \beta_i \mathbf{v}_i}{\left(\sum_{j=1}^{n-1} \frac{1}{|\lambda_1|^2} |\lambda_j \beta_j|^2\right)^{1/2}}. \tag{2}$$

In order to determine what happens in the limit, we need to iterate this expression. To simplify notation, we define $x^{(1)}$ and $y^{(1)}$ as the $\tilde{x}$ and $\tilde{y}$ from the first step of the iteration above. To figure out the next step of the procedure, we note that the eigenvector decompositions of $x^{(1)}$ and $y^{(1)}$ are given by

$$x^{(1)} = \sum_{i=1}^{n-1} \frac{\frac{\lambda_i}{|\lambda_1|} \alpha_i}{\left(\sum_{j=1}^{n-1} \frac{1}{|\lambda_1|^2} |\lambda_j \alpha_j|^2\right)^{1/2}} \mathbf{v}_i := \sum_{i=1}^{n-1} \alpha_i^{(1)} \mathbf{v}_i, \qquad y^{(1)} = \sum_{i=1}^{n-1} \frac{\frac{\lambda_i}{|\lambda_1|} \beta_i}{\left(\sum_{j=1}^{n-1} \frac{1}{|\lambda_1|^2} |\lambda_j \beta_j|^2\right)^{1/2}} \mathbf{v}_i := \sum_{i=1}^{n-1} \beta_i^{(1)} \mathbf{v}_i,$$

so that a second iteration of this procedure gives

$$
\begin{aligned}
x^{(2)} &= \frac{\sum_{i=1}^{n-1} \frac{\lambda_i}{|\lambda_1|}\alpha_i^{(1)}\mathbf{v}_i}{\left(\sum_{j=1}^{n-1} \frac{1}{|\lambda_1|^2}|\lambda_j \alpha_j^{(1)}|^2\right)^{1/2}} \\
&= \frac{\sum_{i=1}^{n-1} \frac{\lambda_i}{|\lambda_1|} \frac{\frac{\lambda_i}{|\lambda_1|}\alpha_i}{\left(\sum_{j=1}^{n-1}\frac{1}{|\lambda_1|^2}|\lambda_j\alpha_j|^2\right)^{1/2}}\mathbf{v}_i}{\left(\sum_{j=1}^{n-1}\frac{1}{|\lambda_1|^2}\left|\lambda_j \frac{\frac{\lambda_j}{|\lambda_1|}\alpha_j}{\left(\sum_{k=1}^{n-1}\frac{1}{|\lambda_1|^2}|\lambda_k\alpha_k|^2\right)^{1/2}}\right|^2\right)^{1/2}} \\
&= \frac{\sum_{i=1}^{n-1}\left(\frac{\lambda_i}{|\lambda_1|}\right)^2\alpha_i\mathbf{v}_i}{\left(\sum_{j=1}^{n-1}\left|\left(\frac{\lambda_j}{|\lambda_1|}\right)^2\alpha_j\right|^2\right)^{1/2}} = \sum_{i=1}^{n-1}\left(\frac{\left(\frac{\lambda_i}{|\lambda_1|}\right)^2\alpha_i}{\left(\sum_{j=1}^{n-1}\left|\left(\frac{\lambda_j}{|\lambda_1|}\right)^2\alpha_j\right|^2\right)^{1/2}}\right)\mathbf{v}_i.
\end{aligned}
$$

Thus, by induction, we can see that

$$
x^{(m)} = \sum_{i=1}^{n-1}\left(\frac{\left(\frac{\lambda_i}{|\lambda_1|}\right)^m\alpha_i}{\left(\sum_{j=1}^{n-1}\left|\left(\frac{\lambda_j}{|\lambda_1|}\right)^m\alpha_j\right|^2\right)^{1/2}}\right)\mathbf{v}_i.
$$

By the same arguments from the power method, we note that as $m \to \infty$, all terms except the ones corresponding to $\lambda_1$ and $\lambda_{n-1}$, the ones with $|\lambda_i| = |\lambda_1|$, will go to zero. Therefore, $x^m$ will approach something of the form

$$
x^{(m)} \approx \left(\frac{\left(\frac{\lambda_1}{|\lambda_1|}\right)^m\alpha_1}{\left(\left|\left(\frac{\lambda_1}{|\lambda_1|}\right)^m\alpha_1\right|^2+\left|\left(\frac{\lambda_{n-1}}{|\lambda_1|}\right)^m\alpha_{n-1}\right|^2\right)^{1/2}}\right)\mathbf{v}_1 + \left(\frac{\left(\frac{\lambda_{n-1}}{|\lambda_1|}\right)^m\alpha_{n-1}}{\left(\left|\left(\frac{\lambda_1}{|\lambda_1|}\right)^m\alpha_1\right|^2+\left|\left(\frac{\lambda_{n-1}}{|\lambda_1|}\right)^m\alpha_{n-1}\right|^2\right)^{1/2}}\right)\mathbf{v}_{n-1}.
$$

However, by our work before, we know that $\alpha_{n-1} = \overline{\alpha_1}$, $\mathbf{v}_{n-1} = \overline{\mathbf{v}_1}$ and $\lambda_n - 1 = \overline{\lambda_1}$. Therefore the two absolute values in the denominator are the same, and we get that

$$
\begin{aligned}
x^{(m)} &\approx \left(\frac{\left(\frac{\lambda_1}{|\lambda_1|}\right)^m\alpha_1}{\left(\left|\left(\frac{\lambda_1}{|\lambda_1|}\right)^m\alpha_1\right|^2+\left|\left(\frac{\overline{\lambda_1}}{|\lambda_1|}\right)^m\overline{\alpha_1}\right|^2\right)^{1/2}}\right)\mathbf{v}_1 + \left(\frac{\left(\frac{\overline{\lambda_1}}{|\lambda_1|}\right)^m\overline{\alpha_1}}{\left(\left|\left(\frac{\lambda_1}{|\lambda_1|}\right)^m\alpha_1\right|^2+\left|\left(\frac{\overline{\lambda_1}}{|\lambda_1|}\right)^m\overline{\alpha_1}\right|^2\right)^{1/2}}\right)\overline{\mathbf{v}_1} \\
&\approx 2\Re\left(\frac{\left(\frac{\lambda_1}{|\lambda_1|}\right)^m\alpha_1\mathbf{v}_1}{\sqrt{2}\left|\left(\frac{\lambda_1}{|\lambda_1|}\right)^m\alpha_1\right|}\right) = 2\Re\left(\left(\frac{\lambda_1}{|\lambda_1|}\right)^m\frac{\alpha_1}{\sqrt{2}|\alpha_1|}\mathbf{v}_1\right).
\end{aligned}
$$

With this calculation, we have proved the following:

**Theorem 3.1.** *Let $P$ be a polygon on $n$ vertices. If the fair averaging procedure is iterated on this polygon, where the process is rescaled at each step to keep the 2-norm of the* $\mathbf{x}$ *and* $\mathbf{y}$ *vectors to be* 1*, then the points* *will converge to lying on the ellipse parametrized by*

$$\frac{1}{\sqrt{n}}\left(2\Re\left(\frac{\alpha_1}{\sqrt{2}|\alpha_1|}\right)\cos(t)-2\Im\left(\frac{\alpha_1}{\sqrt{2}|\alpha_1|}\right)\sin(t), 2\Re\left(\frac{\beta_1}{\sqrt{2}|\beta_1|}\right)\cos(t)-2\Im\left(\frac{\beta_1}{\sqrt{2}|\beta_1|}\right)\sin(t)\right), \qquad 0\le t\le 2\pi,$$

*where* $\alpha_1$ *and* $\beta_1$ *are the coefficients of* $\mathbf{v}_1$ *in the expansions*

$$\mathbf{x}=\sum_{k=1}^{n-1}\alpha_k\mathbf{v}_k, \qquad \mathbf{y}=\sum_{k=1}^{n-1}\beta_k\mathbf{v}_k,$$

*in terms of the eigenvectors of the iteration matrix* $A$*.*

However, in this calculation, we note that the magnitude of the two complex numbers used in the ellipse, $\frac{\alpha_1}{\sqrt{2}|\alpha_1|}$ and $\frac{\beta_1}{\sqrt{2}|\beta_1|}$ are the same. Thus, by Corollary 2.4, we know that this ellipse will be tilted an angle of 45 degrees, corroborating the results of [2].

To fully confirm their results, we need to relate this back to what they got for the vectors. As in the power method, we see that as $m$ goes to $\infty$, the only terms that survive are those with $|\lambda_i|=|\lambda_1|$, namely $\lambda_1$ and $\lambda_{n-1}$. Furthermore, since our vectors are orthonormal, we have that

$$\alpha_1=\mathbf{x}^T\overline{\mathbf{v}_1}, \qquad \beta_1=\mathbf{y}^T\overline{\mathbf{v}_1},$$

so the coefficients we need to analyze are

$$\alpha_1'=\frac{\mathbf{x}^T\overline{\mathbf{v}_1}}{|\mathbf{x}^T\overline{\mathbf{v}_1}|}, \qquad \beta_1'=\frac{\mathbf{y}^T\overline{\mathbf{v}_1}}{|\mathbf{y}^T\overline{\mathbf{v}_1}|}.$$

However, since

$$\mathbf{v}_1=\begin{bmatrix}\cos(0)+i\sin(0)\\ \cos\left(\frac{2\pi}{n}\right)+i\sin\left(\frac{2\pi}{n}\right)\\ \vdots\\ \cos\left(\frac{2\pi}{n}\right)+i\sin\left(\frac{2\pi}{n}\right)\end{bmatrix},$$

and $x$ is a real vector, we have that

$$\mathbf{x}^T\overline{\mathbf{v}_1}=x_0(\cos(0))+x_1\left(\cos\left(\frac{2\pi}{n}\right)-i\sin\left(\frac{2\pi}{n}\right)\right)+\cdots+x_{n-1}\left(\cos\left(\frac{2\pi(n-1)}{n}\right)-i\sin\left(\frac{2\pi(n-1)}{n}\right)\right),$$

from which we can see that

$$\begin{aligned}\sqrt{n}\Re(\mathbf{x}^T\overline{\mathbf{v}_1})&=x_0\cos(0)+x_1\cos\left(\frac{2\pi}{n}\right)+\cdots+x_{n-1}\cos\left(\frac{2\pi(n-1)}{n}\right)\\ \sqrt{n}\Im(\mathbf{x}^T\overline{\mathbf{v}_1})&=-\left(x_0\sin(0)+x_1\sin\left(\frac{2\pi}{n}\right)+\cdots+x_{n-1}\sin\left(\frac{2\pi(n-1)}{n}\right)\right)\\ |\mathbf{x}^T\overline{\mathbf{v}_1}|^2&=\Re(\mathbf{x}^T\overline{\mathbf{v}_1})^2+\Im(\mathbf{x}^T\overline{\mathbf{v}_1})^2.\end{aligned}$$

To match with the notation in [2], we have that $c^T\mathbf{x}^{(0)}=\Re(\mathbf{x}^t\mathbf{v}_1)$ and $s^T\mathbf{x}^{(0)}=-\Im(\mathbf{x}^T\mathbf{v}_1)$. Therefore, this result we see here is exactly the same as what was concluded in the previous paper.

## 3.2 Eigenvalue Normalization

In this paper, we present a new method of normalization for the iterated polygon problem, one that still allows the limiting ellipse to be visualized, but allowing the process to achieve any orientation of the final ellipse. To do this, we consider the same averaging process as before

$$\mathbf{x}^{(m)}=A\mathbf{x}^{(m-1)}, \qquad \mathbf{y}^{(m)}=A\mathbf{y}^{(m-1)},$$

but instead of renormalizing the vectors $\mathbf{x}^{(m)}$ and $\mathbf{y}^{(m)}$ to have norm 1, we instead multiply them by $\frac{1}{|\lambda_1|}$ where $|\lambda_1|$ is the second largest eigenvalue of $A$ in magnitude. If we look back to the power method calculations from the first section, we see that this iterative procedure will result in a limit as $m \to \infty$ of

$$A^m\mathbf{x} = \left(\left(\frac{\lambda_1}{|\lambda_1|}\right)^m \alpha_1\mathbf{v}_1 + \left(\frac{\lambda_{n-1}}{|\lambda_{n-1}|}\right)^m \alpha_{n-1}\mathbf{v}_{n-1} + O\left(\frac{|\lambda_2|}{|\lambda_1|}^m\right)\right),$$

and similarly for $y$. Therefore, by following through the rest of the results in the previous section, we have proved the following:

**Theorem 3.2.** *Let $P$ be a polygon on $n$ vertices. If the fair averaging procedure is iterated on this polygon, where the process is rescaled at each step by multiplying by the inverse magnitude of the second largest eigenvalue, then the points will converge to lying on the ellipse parametrized by*

$$\frac{1}{\sqrt{n}}(2\Re(\alpha_1)\cos(t) - 2\Im(\alpha_1)\sin(t), 2\Re(\beta_1)\cos(t) - 2\Im(\beta_1)\sin(t)), \qquad 0 \le t \le 2\pi,$$

*where $\alpha_1$ and $\beta_1$ are the coefficients of $\mathbf{v}_1$ in the expansions*

$$\mathbf{x} = \sum_{k=1}^{n-1} \alpha_k\mathbf{v}_k, \qquad \mathbf{y} = \sum_{k=1}^{n-1} \beta_k\mathbf{v}_k,$$

*in terms of the eigenvectors of the iteration matrix $A$.*

# 4 Other Extensions

In this section, we look at a generalization of this averaging procedure, where instead of taking the fair average of two consecutive points, we take a weighted average of any number of consecutive points on the polygon.

**Definition 4.1.** Let $n$ be a natural number, and pick $\eta_0, ..., \eta_{n-1}$ real numbers so that

$$\eta_i \ge 0 \ \forall i \qquad \sum_{i=0}^{n-1} \eta_i = 1$$

Then, we define the $\eta$-averaging process on a polygon $P$ with $n$ vertices by, for each vertex $(x_i, y_i)$, the next vertex $(\tilde{x}_i, \tilde{y}_i)$ is given by

$$\tilde{x}_i = \eta_0 x_i + \eta_1 x_{i+1} + \cdots + \eta_{n-1} x_{i+n},$$

and similarly for $\tilde{y}$, where all of the addition in subscripts is taken modulo $n$, i.e., $x_{n+1}$ refers to $x_1$ etc.

For a given vector $\eta$ of values in the definition, we can define an iteration matrix $A_\eta$ for this process as

$$A_\eta = \begin{bmatrix} \eta_0 & \eta_1 & \eta_2 & \eta_3 & \cdots & \eta_n \\ \eta_n & \eta_0 & \eta_1 & \eta_2 & \cdots & \eta_{n-1} \\ \eta_{n-1} & \eta_n & \eta_0 & \eta_1 & \cdots & \eta_{n-2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \eta_1 & \eta_2 & \eta_3 & \eta_4 & \cdots & \eta_0 \end{bmatrix},$$

and as before, the new set of vertices of the polygon is given by $\tilde{\mathbf{x}} = A_\eta\mathbf{x}$ and $\tilde{\mathbf{y}} = A_\eta\mathbf{y}$. From this, we can easily see that the vector $(1/2, 1/2, 0, 0, ..., 0)$ corresponds to the original averaging process. As before, from the eigenvalue calculations, we can write

$$A_\eta = \eta_0 I + \eta_1 S + \cdots + \eta_{n-1} S^{n-1}.$$

**Lemma 4.1.** *The eigenvectors of $A_\eta$ are the set $\{\mathbf{v}_i\}$, the same as the eigenvectors of $S$, and the corresponding eigenvalues are given*

$$\lambda_k = \eta_0 + \eta_1\omega^k + \eta_2\omega^{2k} + \cdots + \eta_{n-1}\omega^{(n-1)k}.$$

**Theorem 4.1.** *Given an averaging vector $\eta$ the corresponding procedure, with proper rescaling, converges to an ellipse in the same manner as the standard averaging procedure.*

*Proof.* Given the conditions of the problem and the restrictions placed on the vector $\eta$, $(1, 1, 1, ..., 1)$ is still an eigenvector of the matrix $A_\eta$ with eigenvalue 1, and all of the remaining eigenvalues have magnitude strictly less than 1. If we assume that there exists a $k$ so that

$$1 > |\lambda_k| = |\lambda_{n-k}| > |\lambda_j| \qquad \forall j \neq 0, k, n-k,$$

then the analysis from the power method carries through exactly as in the previous part. If we have this leading eigenvalue, then we see that, as before, the points converge to lie on the ellipse defined by the corresponding eigenvector. □

## 4.1 Weighted Averages

The simplest generalization of the standard procedure is taking a weighted average of two points on the polygon. For this, we let the vector $\eta = (\xi, 1-\xi, 0, 0, ..., 0)$ so that the matrix $A_\eta$ is given by

$$A_\eta = \begin{bmatrix} \xi & 1-\xi & 0 & 0 & \cdots & 0 \\ 0 & \xi & 1-\xi & 0 & \cdots & 0 \\ 0 & 0 & \xi & 1-\xi & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1-\xi & 0 & 0 & 0 & \cdots & \xi \end{bmatrix}.$$

In this case, we know that the eigenvectors are the same $\{\mathbf{v}_j\}$ that we had for the fair averaging procedure, and the corresponding eigenvalues are

$$\lambda_k = \xi + (1-\xi)\omega^k.$$

By the same arguments as before, we know that these eigenvalues lie on a circle of radius $(1-\xi)$ in the complex plane centered at $\xi$, so that $\lambda_1$ and $\lambda_{n-1}$ are again the largest eigenvalues in magnitude outside of $\lambda_0$, so all of the same calculations and results follow through exactly. We then have the equivalent of both earlier theorems:

**Theorem 4.2.** *Let $P$ be a polygon on $n$ vertices, and let $0 < \xi < 1$ be any real number. Define the $\xi$-averaging process by applying the matrix $A_\eta$ to the $x$ and $y$ components of the vertices of the polygons, where $\eta = (\xi, 1-\xi, 0, ..., 0)$. If this averaging procedure is iterated on this polygon, where the process is rescaled at each step to keep the 2-norm of the $\mathbf{x}$ and $\mathbf{y}$ vectors to be 1, then the points will converge to lying on the ellipse parametrized by*

$$\frac{1}{\sqrt{n}}\left(2\Re\left(\frac{\alpha_1}{\sqrt{2}|\alpha_1|}\right)\cos(t) - 2\Im\left(\frac{\alpha_1}{\sqrt{2}|\alpha_1|}\right)\sin(t), 2\Re\left(\frac{\beta_1}{\sqrt{2}|\beta_1|}\right)\cos(t) - 2\Im\left(\frac{\beta_1}{\sqrt{2}|\beta_1|}\right)\sin(t)\right), \qquad 0 \leq t \leq 2\pi,$$

*where $\alpha_1$ and $\beta_1$ are the coefficients of $\mathbf{v}_1$ in the expansions*

$$\mathbf{x} = \sum_{k=1}^{n-1}\alpha_k\mathbf{v}_k \qquad \mathbf{y} = \sum_{k=1}^{n-1}\beta_k\mathbf{v}_k$$

*in terms of the eigenvectors of the iteration matrix $A$.*

**Theorem 4.3.** *Let $P$ be a polygon on $n$ vertices, and let $0 < \xi < 1$ be any real number. Define the $\xi$-averaging process by applying the matrix $A_\eta$ to the $x$ and $y$ components of the vertices of the polygons, where $\eta = (\xi, 1-\xi, 0, ..., 0)$. If this averaging procedure is iterated on this polygon, where the process is rescaled at each step by multiplying by the inverse magnitude of the second largest eigenvalue, then the points will converge to lying on the ellipse parametrized by*

$$\frac{1}{\sqrt{n}}(2\Re(\alpha_1)\cos(t) - 2\Im(\alpha_1)\sin(t), 2\Re(\beta_1)\cos(t) - 2\Im(\beta_1)\sin(t)), \qquad 0 \le t \le 2\pi,$$

*where $\alpha_1$ and $\beta_1$ are the coefficients of $\mathbf{v}_1$ in the expansions*

$$\mathbf{x} = \sum_{k=1}^{n-1} \alpha_k \mathbf{v}_k \qquad \mathbf{y} = \sum_{k=1}^{n-1} \beta_k \mathbf{v}_k$$

*in terms of the eigenvectors of the iteration matrix $A$.*

# 5 Maple Packages

This note includes three Maple packages to illustrate the numerics of this method. The first, PolygonHelp, contains the documentation for the other two files. The second, PolygonSupport, contains the supporting methods necessary to run the experiment, implementing some linear algebra helper methods for the iteration and analysis. The last, PolygonProject, which reads the PolygonSupport package, implements the iteration procedures presented in Sections 2 and 3 and displays the results. The main procedures to run to illustrate the results here are `AnimateConvergence` and `EllipseProperties`, with `RandomGoodPoly` as the way to generate a polygon to initialize the process. See the website

`http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/polygon.html`

for access to the packages.


# Acknowledgements

We would like to thank the other students in the Experimental Math Class during Spring 2018, including Ahsan Khan and Edna Jones, who worked out some of the initial details from this paper during the semester.


# References


[1] Kendall Atkinson. *An Introduction to Numerical Analysis.* Wiley, second edition, 1989.

[2] Adam N. Elmachtoub and Charles F. Van Loan. From random polygon to ellipse: An eigenanalysis. *SIAM Review*, 52(1):151–170, 2100.

[3] Charles F. Van Loan. Untangling random polygons and other things. *SIAM News*, 51, 2018.