involve
a journal of mathematics

# Convergence of sequences of polygons

Eric Hintikka and Xingping Sun

msp

2016 vol. 9, no. 5

INVOLVE 9:5 (2016)
dx.doi.org/10.2140/involve.2016.9.751

# Convergence of sequences of polygons

Eric Hintikka and Xingping Sun

(Communicated by Zuhair Nashed)

In 1878, Darboux studied the problem of midpoint iteration of polygons. Simply put, he constructed a sequence of polygons $\Pi^{(0)}, \Pi^{(1)}, \Pi^{(2)}, \ldots$ in which the vertices of a descendant polygon $\Pi^{(k)}$ are the midpoints of its parent polygon $\Pi^{(k-1)}$ and are connected by edges in the same order as those of $\Pi^{(k-1)}$. He showed that such a sequence of polygons converges to their common centroid. In proving this result, Darboux utilized the powerful mathematical tool we know today as the finite Fourier transform. For a long time period, however, neither Darboux's result nor his method was widely known. The same problem was proposed in 1932 by Rosenman as Monthly Problem # 3547 and had been studied by several authors, including I. J. Schoenberg (1950), who also employed the finite Fourier transform technique. In this paper, we study generalizations of this problem. Our scheme for the construction of a polygon sequence not only gives freedom in selecting the vertices of a descendant polygon but also allows the polygon generating procedure itself to vary from one step to another. We show under some mild restrictions that a sequence of polygons thus constructed converges to a single point. Our main mathematical tools are ergodicity coefficients and the Perron–Frobenius theory on nonnegative matrices.

## 1. Introduction

Jean Gaston Darboux [1878] proposed and solved the following problem. Let $\Pi^{(0)}$ be a closed polygon in the plane with vertices

$$v_0^{(0)}, v_1^{(0)}, \ldots, v_{n-1}^{(0)}.$$

Denote by

$$v_0^{(1)}, v_1^{(1)}, \ldots, v_{n-1}^{(1)},$$

respectively, the midpoints of the edges $v_0^{(0)}v_1^{(0)},\ v_1^{(0)}v_2^{(0)}, \ldots, v_{n-1}^{(0)}v_0^{(0)}$. Connecting $v_0^{(1)}, v_1^{(1)}, \ldots, v_{n-1}^{(1)}$ in the same order as above, we derive a new polygon, denoted by $\Pi^{(1)}$. Apply the same procedure to derive polygon $\Pi^{(2)}$. After $k$ constructions, we obtain polygon $\Pi^{(k)}$. Show that $\Pi^{(k)}$ converges, as $k \to \infty$, to the centroid of the

*MSC2010:* primary 15A60, 65F35; secondary 15A12, 15A30.
*Keywords:* polygons, finite Fourier transform, stochastic matrices, ergodicity coefficients.

original points $v_0^{(0)}, v_1^{(0)}, \ldots, v_{n-1}^{(0)}$. We will refer to this problem as "midpoint iteration of polygons". For any given sequence of polygons $\Pi^{(0)}, \Pi^{(1)}, \Pi^{(2)}, \ldots$, we will call $\Pi^{(k)}$ the descendant polygon of $\Pi^{(k-1)}$ and $\Pi^{(k-1)}$ the parent polygon of $\Pi^{(k)}$.

In proving his result, Darboux used the powerful mathematical tool we know today as the finite Fourier transform. This allowed him to establish an exponential rate at which the polygon sequence converges to their common centroid. For a long period of time, however, neither Darboux's result nor his method was widely known. More than half a century later, the same problem, which has since been known as Monthly Problem # 3547, was proposed by Rosenman, and a solution of the problem was given by R. Huston in [Rosenman and Huston 1933].

Unaware of what Darboux had already done, Schoenberg [1950] completely retooled the finite Fourier transform technique to tackle the problem of midpoint iteration of polygons. Schoenberg also generalized the problem by allowing vertices of a descendant polygon to come from convex hulls of consecutive vertices of its parent polygon. Later, Schoenberg [1982] revisited this interesting topic. Terras [1999] summarized Schoenberg's work as an example of applications of the finite Fourier transform. One can approach the problem of midpoint iteration of polygons from other mathematical perspectives. For example, Ding et al. [2003] and Ouyang [2013] considered this problem as a special case of Markov chains, and Treatman and Wickham [2000] studied a logarithmic dual problem in which all the vertices of the polygons are on the unit circle and the convergence is to a regular polygon.

In this paper, we study several generalizations of this problem. In Section 3, we consider cases in which the vertices of a descendant polygon are not necessarily midpoints of the edges of its parent polygon but can be chosen more freely from the edges of its parent polygon. In Section 4, we further generalize the work done in Section 3 by allowing the polygon generating procedure to vary from one step to another. In Section 5, we again elevate the level of freedom in selecting the vertices of a descendant polygon by allowing them to come from convex hulls of some subsets of the vertices of its parent polygon. Technically, Section 4 deals with a special case of what is studied in Section 5. In our opinion, however, the importance of the special case deserves some special attention, as does the mathematical argument employed therein. Furthermore, our results in Section 4 are more quantitative, and their geometric implications more illustrative. In addition, the flow of representation reflects the progressive nature of our research process. Section 2 is devoted to the introduction of frequently used notations and definitions.

To conclude the introduction of this paper, we share with readers a few highlights of this research experience. In the midpoint polygon iteration problem, if we view the collection of vertices of a polygon as a vector $z := (z_0, z_1, \ldots, z_{n-1})^T$ in $\mathbb{C}^n$, then the collection of vertices of the first descendant polygon is $Az$, where

$$A := \operatorname{circ}\left(\tfrac{1}{2}, \tfrac{1}{2}, 0, \dots, 0\right),$$

in which $\operatorname{circ}\left(\tfrac{1}{2}, \tfrac{1}{2}, 0, \dots, 0\right)$ denotes the $n \times n$ circulant matrix[1] whose first row is $\left(\tfrac{1}{2}, \tfrac{1}{2}, 0, \dots, 0\right)$. Likewise, the $k$-th polygon has vertices $A^k z$. Scrutinizing Darboux and Schoenberg's proofs, we found that they had implicitly established the stronger result that $\|A^k - L\|_2$ converges to zero exponentially, where $\|\cdot\|_2$ indicates the spectral radius norm for square matrices, and $L$ is the rank-one matrix whose entries are all $1/n$. It follows that the sequence of the polygons converges to their common centroid. We briefly entertained several possible ways to generalize this problem before we chose to focus on investigating the asymptotic behavior of products of (square) row stochastic matrices and the geometric implications for the corresponding sequence of polygons. Witnessing that the finite Fourier transform works wonderfully with circulant matrices, we tried bounding an arbitrary stochastic matrix by a sum of circulant stochastic matrices. While we have had some success with this strategy in estimating the smallest eigenvalue of a nonsingular stochastic matrix, we have yet to retool the method in a suitable way for the problem in this paper. Our basic tools in this paper are the Perron–Frobenius theorem [Horn and Johnson 1990] on nonnegative matrices and ergodicity coefficients [Ipsen and Selee 2011].

## 2. Notations and definitions

We will use boldface letters, such as $\boldsymbol{v}$, to denote vectors in $\mathbb{C}^n$. The $i$-th component of $\boldsymbol{v}$ is denoted by $v_i$. When the full form of the vector $\boldsymbol{v}$ is needed in some context, we will write $\boldsymbol{v} = (v_0, v_1, \dots, v_{n-1})^T$.

Let $n$ complex numbers (not necessarily distinct) be given. We may connect them in any given order to form a (possibly degenerate) $n$-gon in the complex plane. In this way, an $n$-gon can be identified with a vector in $\mathbb{C}^n$, and vice versa. Label the $n$ complex numbers according to the order in which they are connected by edges: $v_0, v_1, \dots, v_{n-1}, v_n, \dots$. That is, two components are adjacent if and only if the corresponding vertices are connected by an edge. To facilitate mathematical exposition, we have here adopted arithmetic modulo $n$. For example, $v_0$ and $v_n$ are the same vertex. We will use the same modular arithmetic for row and column indices of matrix entries, announcing as we do so.

If $A$ is a matrix, we denote by $(A)_{ij}$ the entry of $A$ located at the $i$-th row and the $j$-th column. If $A$ is a square matrix, then the *spectral radius* of $A$ is denoted by $\rho(A)$, and we define

$$\rho(A) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}.$$

[1] The definition of circulant matrices will be given in Section 2.

**Definition 1.** A matrix $A$ is *positive*, denoted by $A > 0$, if $(A)_{ij} > 0$ for all $i, j$. Similarly, $A$ is *nonnegative*, or $A \geq 0$, if $(A)_{ij} \geq 0$ for all $i, j$. If, for some $\alpha \in \mathbb{R}$, we have $(A)_{ij} > \alpha$ (respectively, $(A)_{ij} \geq \alpha$) for all $i, j$, then we will write $A > \alpha$ (respectively, $A \geq \alpha$).

**Definition 2.** A *stochastic* (or *row-stochastic*) matrix is a real-valued, nonnegative, square matrix whose row sums are all 1.

**Definition 3.** An $n \times n$ matrix $A$ is *circulant* if for some complex numbers $a_i$, we have

$$A = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-3} & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_2 & a_3 & \cdots & a_0 & a_1 \\ a_1 & a_2 & \cdots & a_{n-1} & a_0 \end{bmatrix}.$$

In Section 1, we used the notation $\operatorname{circ}(a_0, a_1, \ldots, a_{n-2}, a_{n-1})$ to denote the above circulant matrix. We will continue to do so in appropriate contexts.

**Definition 4.** We call an $n \times m$ matrix $A$ *$k$-banded* if

$$(A)_{ij} \neq 0 \iff j \in \{i, i+1, \ldots, i+k-1\} \pmod n.$$

For example, the matrix $\operatorname{circ}(1/2, 1/2, 0, \ldots, 0)$ is a 2-banded matrix.

**Definition 5.** We say that two $n \times m$ matrices $A$ and $B$ have the same *zero pattern* if $(A)_{ij} = 0 \iff (B)_{ij} = 0$ for all $i$, $j$.

**Definition 6.** We say that an $n \times n$ matrix is *circulant-patterned* if it has the same zero pattern as a circulant matrix.

**Definition 7.** We say that a sequence of $n$-gons $(\Pi^{(k)})_{k \geq 0}$ *converges to a point $q \in \mathbb{C}$* if, for any $1 \leq p \leq \infty$, we have $\lim_{k \to \infty} \|\Pi^{(k)} - \boldsymbol{q}\|_p = 0$, where $\boldsymbol{q} = (q, q, \ldots, q)^T$. Here $\|\cdot\|_p$ denotes the $p$-norm on $\mathbb{C}^n$, that is,

$$\|\boldsymbol{v}\|_p = \begin{cases} \left(\sum_{i=1}^n |v_i|^p\right)^{1/p} & \text{if } 1 \leq p < \infty, \\ \max\limits_{1 \leq i \leq n} |v_i| & \text{if } p = \infty. \end{cases}$$

In this paper, we primarily work with the 1-norm. To be sure, any two norms on a finite-dimensional normed linear space are topologically equivalent.

## 3. Polygons derived from a fixed 2-banded matrix

In this section, we suppose that $\Pi^{(0)}$ is an $n$-gon and that its $k$-th descendant polygon is given by $\Pi^{(k)} = A^k \Pi^{(0)}$, where $A$ is a fixed 2-banded stochastic matrix. Hence, for some real $\alpha_0, \alpha_1, \dots, \alpha_{n-1}$ with $0 < \alpha_i < 1$, we have

$$A = \begin{bmatrix} \alpha_0 & 1-\alpha_0 & 0 & \cdots & 0 & 0 \\ 0 & \alpha_1 & 1-\alpha_1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & & \vdots \\ 1-\alpha_{n-1} & 0 & \cdots & 0 & 0 & \alpha_{n-1} \end{bmatrix}. \tag{3-1}$$

We are interested in this particular construction because, geometrically speaking, the vertices of $\Pi^{(k+1)}$ are chosen from within the edges of $\Pi^{(k)}$, one vertex per edge. The procedure also allows the choice of any particular vertex in $\Pi^{(k+1)}$ to be independent from the others. We aim to show that $(\Pi^{(k)})_{k\geq 0}$ converges to a predetermined point. In the present section, the theoretic foundation for our argument is Perron's theorem (8.2.11(f) in [Horn and Johnson 1990]), which we state in the following theorem. To be sure, the convergence results in this section follow from the general framework of the Perron–Frobenius theorem. However, the 2-banded structure of our matrices allows us to obtain more nuanced convergence results. In particular, our knowledge of the convergence process is quantitative in the sense that we are able to predetermine the point to which the sequence of polygons converges.

**Theorem** (Perron). *If $A$ is a positive $n \times n$ matrix, then*

$$[\rho(A)^{-1}A]^m \to L \quad \textit{as } m \to \infty,$$

*where $L = xy^T$, $Ax = \rho(A)x$, $A^Ty = \rho(A)y$, $x > 0$, $y > 0$, and $x^Ty = 1$.*

We now derive some quick results and use these, along with Perron's theorem, to show that the sequence $(\Pi^{(k)})_{k\geq 0}$ in fact converges to a point for any choice of $A$. Additionally, we give an expression for that limiting point in terms of the entries of $A$.

**Proposition 8.** *If $A$ and $A_i$, where $i \in \{0, 1, \dots, k-1\}$, are $n \times n$ stochastic matrices, then we have:*

(1) *The spectral radius $\rho(A)$ is* 1.

(2) *The product matrix $A_{k-1}A_{k-2}\cdots A_0$ is stochastic.*

*Proof.* These are known results. Part (1) follows from Lemma 8.1.21 in [Horn and Johnson 1990]. We give a short yet entertaining proof to part (2) using the simple fact that an $n \times n$ matrix $A$ is stochastic if and only if $Ae = e$, where $e \in \mathbb{R}^n$ is the vector with all components 1. We simply write

$$A_{k-1}A_{k-2}\cdots A_0 e = A_{k-1}A_{k-2}\cdots A_1 e = \cdots = e. \qquad \square$$

**Proposition 9.** *Suppose that $A$ is a* 2*-banded stochastic matrix as given in* (3-1). *Then $A^{n-1} > 0$.*

This result is stated in [Ouyang 2013] without a proof. We give a complete proof here, as variations of it will become quite useful in the latter part of the paper.

*Proof.* Throughout the proof, we use arithmetic modulo $n$ to track the changes in row and column indices as a result of matrix multiplications. For $n = 2$, the result is obvious. Suppose that for some $N \in \mathbb{N}$, we have $(A^N)_{ij} > 0$ for all $i$ and $j$ such that $j \in \{i, i+1, \dots, i+N\} \pmod n$. Then for any such $i$ and $j$, we have

$$(A^{N+1})_{ij} = \sum_{k=0}^{n-1} (A^N)_{ik}(A)_{kj} \geq (A^N)_{ij}(A)_{jj} > 0.$$

Furthermore, we have

$$\begin{aligned}(A^{N+1})_{i,i+N+1} &= \sum_{k=0}^{n-1} (A)_{ik}(A^N)_{k,i+N+1} \\ &= \alpha_i (A^N)_{i,i+N+1} + (1-\alpha_i)(A^N)_{i+1,i+N+1} \\ &\geq (1-\alpha_i)(A^N)_{i+1,i+N+1},\end{aligned}$$

which is positive by the induction hypothesis. It follows that the matrix $A^{N+1}$ has positive entries at $(i, j)$ whenever $j \in \{i, i+1, \dots, i+N+1\} \pmod n$. Hence $A^{n-1}$ has positive entries everywhere. □

**Proposition 10.** *Let $A$ be a matrix as given in* (3-1). *Then we have*

$$\lim_{k \to \infty} A^k = L,$$

*where $L$ is the rank-one matrix given by* (3-2) *in the proof below.*

*Proof.* Let $B = A^{n-1}$. Since $B$ is the product of $(n-1)$ stochastic matrices, it is itself stochastic by Proposition 8. Furthermore, we have that $\rho(B) = 1$. Let

$$y = ((1-\alpha_0)^{-1}, (1-\alpha_1)^{-1}, \dots, (1-\alpha_{n-1})^{-1})^T.$$

One can verify that $A^T y = y$. Thus, $B^T y = (A^{n-1})^T y = (A^T)^{n-1} y = y = \rho(B) y$. Let $x = \alpha_A\,(1, \dots, 1)^T$, where $\alpha_A$ is the scalar given by $\alpha_A = \bigl(\sum_{i=0}^{n-1} (1-\alpha_i)^{-1}\bigr)^{-1}$. Then we have that $Bx = \rho(B)x$ and that $x^T y = 1$. Let $L = xy^T$. Since $x > 0$ and $y > 0$, the rank-one matrix $L$ has identical rows. Specifically,

$$L = \begin{bmatrix} \Bigl((1-\alpha_0)\sum_{i=0}^{n-1} \frac{1}{1-\alpha_i}\Bigr)^{-1} & \cdots & \Bigl((1-\alpha_{n-1})\sum_{i=0}^{n-1} \frac{1}{1-\alpha_i}\Bigr)^{-1} \\ \vdots & & \vdots \\ \Bigl((1-\alpha_0)\sum_{i=0}^{n-1} \frac{1}{1-\alpha_i}\Bigr)^{-1} & \cdots & \Bigl((1-\alpha_{n-1})\sum_{i=0}^{n-1} \frac{1}{1-\alpha_i}\Bigr)^{-1} \end{bmatrix}. \tag{3-2}$$

By Proposition 9, we have $B > 0$. Applying Perron's theorem, we conclude that

$$\lim_{k \to \infty} B^k = L.$$

The rest of the proof is devoted to showing that $\lim_{k\to\infty} A^k = L$. Since $L$ has identical rows and $A^i$ is a stochastic matrix for any $i \in \{0, 1, \dots, n-2\}$, we have, for all $l, m$, that

$$(A^i L)_{lm} = \sum_{k=0}^{n-1} (A^i)_{lk} (L)_{km} = \sum_{k=0}^{n-1} (A^i)_{lk} (L)_{lm} = (L)_{lm} \sum_{k=0}^{n-1} (A^i)_{lk} = (L)_{lm},$$

that is, $L = A^i L$. Hence we have

$$L = A^i \lim_{k\to\infty} B^k = \lim_{k\to\infty} A^i A^{k(n-1)} = \lim_{k\to\infty} A^{k(n-1)+i} \quad \text{for } 0 \le i \le n-2.$$

For a given $\epsilon > 0$, let $N_i$ be such that $\|A^{m(n-1)+i} - L\| < \epsilon$ for all $m \ge N_i$. Let $N = \max\{N_i : i \in \{0, 1, \dots, n-2\}\}$. Choose $j \ge N(n-1)+(n-2)$. By the division theorem, $j = m(n-1)+i$ for some integer $m$ and some fixed $i \in \{0, 1, \dots, n-2\}$. So,

$$j = m(n-1) + i \ge N(n-1) + (n-2) \ge N(n-1) + i \ge N_i(n-1) + i.$$

Hence $m \ge N_i$. Thus we have $\|A^j - L\| < \epsilon$. This inequality holds true for all $j \ge N(n-1) + (n-2)$, which proves that $\lim_{k\to\infty} A^k = L$. □

As the main theorem of this section, we restate the result of Proposition 10 in terms of convergence of a sequence of polygons.

**Theorem 11.** *Let* $(\Pi^{(k)})_{k\ge 0}$ *be a polygon sequence constructed by* $\Pi^{(k)} = A^k \Pi^{(0)}$, *where $A$ is given as in* (3-1). *Then we have*

$$\lim_{k\to\infty} \Pi^{(k)} = (q, \dots, q)^T,$$

*where*

$$q = \sum_{j=0}^{n-1} \left( (1-\alpha_j) \sum_{i=0}^{n-1} \frac{1}{1-\alpha_i} \right)^{-1} \Pi_j^{(0)}.$$

We remind readers that a matrix $A$ given as in (3-1) is circulant if and only if $\alpha_i = \alpha_j$ for all $i, j$. When this holds true, we have

$$(q, \dots, q)^T = \tfrac{1}{n}(\Pi_0^{(0)} + \Pi_1^{(0)} + \cdots + \Pi_{n-1}^{(0)}),$$

which is the centroid of the vertices of $\Pi^{(0)}$. The special case that $\alpha_i = \frac{1}{2}$ for all $i$ corresponds to the problem of midpoint iteration of polygons.

## 4. Polygons derived from a sequence of 2-banded matrices

Let $\delta \in \left(0, \frac{1}{2}\right)$. For each $k \in \mathbb{N}$, we arbitrarily choose $n$ numbers $\alpha_0^{(k)}, \alpha_1^{(k)}, \dots, \alpha_{n-1}^{(k)}$ from the open interval $(\delta, 1-\delta)$ and form the matrix

$$A_k = \begin{bmatrix} \alpha_0^{(k)} & 1-\alpha_0^{(k)} & 0 & \cdots & 0 & 0 \\ 0 & \alpha_1^{(k)} & 1-\alpha_1^{(k)} & 0 & \cdots & 0 \\ \vdots & & \ddots & & & \vdots \\ 1-\alpha_{n-1}^{(k)} & 0 & \cdots & 0 & 0 & \alpha_{n-1}^{(k)} \end{bmatrix}. \tag{4-1}$$

Let $\mathcal{A}_k = A_k A_{k-1} \cdots A_1$. Let $\Pi^{(0)}$ be an $n$-gon, and define $\Pi^{(k)} = \mathcal{A}_k \Pi^{(0)}$. We will show that the sequence of polygons $(\Pi^{(k)})_{k\geq 1}$ converges to a point. Under these circumstances, Perron's theorem is no longer applicable. Our argument relies on some key properties of ergodicity coefficients thoroughly studied in a recent article by Ipsen and Selee [2011].

**Definition 12.** The 1-norm ergodicity coefficient $\tau_1(S)$ for an $n \times n$ stochastic matrix $S$ is given by

$$\tau_1(S) = \max_{\substack{\|z\|_1=1 \\ z^T e=0}} \|S^T z\|_1,$$

where $e = (1, \dots, 1)^T \in \mathbb{R}^n$ and the maximum ranges over $z \in \mathbb{R}^n$. If $n = 1$, we say that $\tau_1(S) = 0$.

**Proposition 13.** *If $S$, $S_1$, and $S_2$ are stochastic matrices, then*:

(1) $0 \leq \tau_1(S) \leq 1$. *Furthermore*, $\tau_1(S) = 0 \Longleftrightarrow S$ *is a rank-one matrix.*

(2) $|\lambda| \leq \tau_1(S)$ *for all eigenvalues* $\lambda < 1$ *of* $S$.

(3) $\tau_1(S) = \frac{1}{2} \max_{i,j} \sum_{k=1}^n |(S)_{ik} - (S)_{jk}|$.

(4) $\tau_1(S_1 S_2) \leq \tau_1(S_1)\tau_1(S_2)$.

*Proof.* Part (1) is from Theorem 3.4 in [Ipsen and Selee 2011], while parts (2) and (4) are the results of Theorem 3.6, and part (3) is the result of Theorem 3.7, of the same work. □

Ergodicity coefficients can be defined and studied for all $p$-norms and even more general metrics under broad matrix analysis settings. For our purpose, however, the results in Proposition 13 suffice.

To proceed, we need the following generalization of Proposition 9.

**Proposition 14.** *For* $1 \leq k \leq n-1$, *let* $A_k$ *be as defined in* (4-1). *Then*

$$\mathcal{A}_{n-1} > \delta^{n-1}.$$

*Proof.* The proof can be considered as a quantification of that of Proposition 9. Arithmetic modulo $n$ will be used to track the changes in row and column indices stemming from matrix multiplications. Suppose that for some $N \in \{1, 2, \dots, n-2\}$ we have $(\mathcal{A}_N)_{ij} > \delta^N$ for all $i \in \{1, 2, \dots, n\}$ and $j \in \{i, i+1, \dots, i+N\}$. Then, for all such $i$ and $j$,

$$\begin{aligned}(\mathcal{A}_{N+1})_{ij} &= \sum_{k=1}^n (A_{N+1})_{ik} (\mathcal{A}_N)_{kj} \\ &\geq (A_{N+1})_{ii} (\mathcal{A}_N)_{ij} > \delta \cdot \delta^N = \delta^{N+1},\end{aligned}$$

that is, $(\mathcal{A}_{N+1})_{ij} > \delta^{N+1}$. Also,

$$(\mathcal{A}_{N+1})_{i,(i+N+1)} = \alpha_i^{(N+1)} \cdot (\mathcal{A}_N)_{i,(i+N+1)} + (1-\alpha_i^{(N+1)}) \cdot (\mathcal{A}_N)_{(i+1),(i+N+1)}$$
$$\geq (1-\alpha_i^{(N+1)}) \cdot (\mathcal{A}_N)_{(i+1),(i+N+1)} > \delta \cdot \delta^N = \delta^{N+1}.$$

Thus $(\mathcal{A}_{N+1})_{ij} > \delta^{N+1}$ for all $i \in \{1, 2, \ldots, n\}$ and $j \in \{i, i+1, \ldots, i+N+1\}$.

Since in the case $N = 1$ it is clearly true that $(\mathcal{A}_N)_{ij} > \delta^N$ for all $i \in \{1, 2, \ldots, n\}$ and $j \in \{i, (i \bmod n)+1\}$, it follows from the principle of mathematical induction that $(\mathcal{A}_{n-1})_{ij} > \delta^{n-1}$ for all $i, j \in \{1, 2, \ldots, n\}$. □

**Proposition 15.** *If $S$ is a positive $n \times n$ stochastic matrix and*

$$\epsilon := \min_{i,j} (S)_{ij},$$

*then*

$$\tau_1(S) \leq 1 - n\epsilon.$$

*Proof.* We first point out that under the conditions specified in Proposition 15, we have $n\epsilon \leq 1$. Therefore, the number on the right-hand side of the above inequality is nonnegative. Let $S_0$ be the $n \times n$ matrix defined by $(S_0)_{ij} = (S)_{ij} - \epsilon$ for all $i, j$. Then $S_0$ is nonnegative, and the row sums of $S_0$ are all $1 - n\epsilon$. More pertinently, we have $\tau_1(S) = \tau_1(S_0)$. To calculate $\tau_1(S_0)$, we write, for all $i, j \in \{1, 2, \ldots, n\}$, that

$$\sum_{k=1}^{n} |(S_0)_{ik} - (S_0)_{jk}| \leq \sum_{k=1}^{n} |(S_0)_{ik}| + \sum_{k=1}^{n} |(S_0)_{jk}| \leq 2(1 - n\epsilon).$$

The desired result then follows from part (3) of Proposition 15. □

We state our main result of this section in the following theorem.

**Theorem 16.** *Let $A_\ell$ $(0 \leq \ell < \infty)$ be a sequence of matrices as given in* (4-1), *and let $\mathcal{A}_k = A_0 A_1 \cdots A_k$. Then we have*

$$\lim_{k\to\infty} \mathcal{A}_k = L,$$

*where $L$ is a rank-one stochastic matrix with identical rows. Hence if $\Pi^{(k)}$ is the corresponding sequence of polygons, we have*

$$\lim_{k\to\infty} \Pi^{(k)} = L\Pi^{(0)},$$

*and thus $(\Pi^{(k)})_{k\geq 0}$ converges to a point.*

*Proof.* Let $(A_\ell)_{\ell\geq 0}$ be matrices as given in (4-1). For each $k \geq 1$, define

$$\mathcal{B}_k = A_{kn-1} A_{kn-2} \cdots A_{(k-1)n}.$$

Then by Propositions 14 and 15, we have

$$\tau_1(\mathcal{B}_k) \leq 1 - n\delta^{n-1} \quad \text{for } k \geq 1.$$

For a given $k \geq 1$, let $m = \max\{j : jn-1 \leq k\}$. By parts (1) and (4) of Proposition 13, we have

$$\begin{aligned}\tau_1(\mathcal{A}_k) &= \tau_1(A_k A_{k-1} \cdots A_{mn} \mathcal{B}_m \mathcal{B}_{m-1} \cdots \mathcal{B}_1)\\ &\leq \tau_1(\mathcal{B}_m \mathcal{B}_{m-1} \cdots \mathcal{B}_1)\\ &\leq \tau_1(\mathcal{B}_m)\tau_1(\mathcal{B}_{m-1}) \cdots \tau_1(\mathcal{B}_1)\\ &\leq (1 - n\delta^{n-1})^m.\end{aligned}$$

Note that $m \to \infty$ when $k$ does. Thus we have

$$\lim_{k\to\infty} \tau_1(\mathcal{A}_k) \leq \lim_{k\to\infty} (1 - n\delta^{n-1})^m = 0.$$

It follows from part (1) of Proposition 13 that $\mathcal{A}_k$ converges to a rank-one matrix. To show that $L$ has identical rows, we use a Cauchy sequence argument. For any given $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $m \geq N$, we have

$$\frac{1}{2} \max_{i,j} \sum_{l=1}^{n} |(\mathcal{A}_m)_{il} - (\mathcal{A}_m)_{jl}| < \frac{\epsilon}{4}.$$

This implies that

$$\max_{i,j} |(\mathcal{A}_m)_{ij} - (\mathcal{A}_m)_{1j}| < \frac{\epsilon}{2}. \tag{4-2}$$

This allows us to write

$$(\mathcal{A}_m)_{ij} = a_j + \delta_{ij}^{(m)} \quad \text{for } 1 \leq i,\, j \leq n,$$

in which $a_j$ is fixed for each $1 \leq j \leq n$, and

$$|\delta_{ij}^{(m)}| \leq \frac{\epsilon}{2} \quad \text{for } 1 \leq j \leq n \text{ and } m > N.$$

Upon writing $\mathcal{A}_{m+k} = S_k \mathcal{A}_m$, where $S_k$ is a stochastic matrix, we have

$$\begin{aligned}(\mathcal{A}_{m+k})_{ij} &= \sum_{l=1}^{n} (S_k)_{il} (\mathcal{A}_m)_{lj} = \sum_{l=1}^{n} (S_k)_{il} (a_j + \delta_{lj}^{(m)})\\ &= a_j \sum_{l=1}^{n} (S_k)_{il} + \sum_{l=1}^{n} (S)_{il} \delta_{lj}^{(m)}\\ &= a_j + \sum_{l=1}^{n} (S_k)_{il} \delta_{lj}^{(m)}.\end{aligned}$$

We also have that

$$-\frac{\epsilon}{2} = -\frac{\epsilon}{2} \sum_{l=1}^{n} (S_k)_{il} < \sum_{l=1}^{n} (S_k)_{il} \delta_{lj}^{(m)} < \frac{\epsilon}{2} \sum_{l=1}^{n} (S_k)_{il} = \frac{\epsilon}{2}.$$

Hence we have $|(\mathcal{A}_{m+k})_{ij} - a_j| < \epsilon/2$, that is,

$$|(\mathcal{A}_{m+k})_{ij} - (\mathcal{A}_m)_{1j}| < \frac{\epsilon}{2}. \tag{4-3}$$

We combine inequalities (4-2) and (4-3) to have

$$\begin{aligned}|(\mathcal{A}_{m+k})_{ij} - (\mathcal{A}_m)_{i'j}| &= |(\mathcal{A}_{m+k})_{ij} - (\mathcal{A}_m)_{1j} + (\mathcal{A}_m)_{1j} - (\mathcal{A}_m)_{i'j}| \\ &\le |(\mathcal{A}_{m+k})_{ij} - (\mathcal{A}_m)_{1j}| + |(\mathcal{A}_m)_{1j} - (\mathcal{A}_m)_{i'j}| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,\end{aligned}$$

which is true for all $m > N$, $k \ge 0$, and all $0 \le i, i', j \le n-1$. The above inequality shows that for each fixed $i$ and $j$, the sequence $(\mathcal{A}_k)_{ij}$ is Cauchy, and that for each fixed $0 \le j \le n-1$, the limits of the sequences $(\mathcal{A}_k)_{ij}$ are the same for all $0 \le i \le n-1$. Thus, for each $0 \le j \le n-1$, there exists a real number $q_j$ such that $\lim_{k\to\infty} (\mathcal{A}_k)_{ij} = q_j$ for all $0 \le i \le n-1$. Hence $\mathcal{A}_k$ converges to the rank-one matrix

$$\begin{bmatrix} q_0 & q_1 & \cdots & q_{n-1} \\ q_0 & q_1 & \cdots & q_{n-1} \\ \vdots & \vdots & & \vdots \\ q_0 & q_1 & \cdots & q_{n-1} \end{bmatrix}. \qquad \square$$

In the above proof, the $ij$-entries of the rank-one matrix $L$ are given as limits of the sequences $(\mathcal{A}_k)_{ij}$. Since they determine the position where the sequence of the polygons converges, a certain effort should be devoted to finding the limits. Doing so, however, would have gone beyond the scope of this paper.

## 5. Polygons derived from a sequence of circulant-patterned matrices

In the previous two sections, we were concerned specifically with polygons derived from sequences of 2-banded stochastic matrices. Each descendant polygon thus generated is inscribed in its parent polygon, a fact which may be utilized to control polygons of other types. In this section, we broaden our scope and consider polygons derived from sequences of matrices of a more general class, namely, stochastic circulant-patterned matrices.

**Proposition 17.** *Suppose that $(A_\ell)_{\ell\ge 0}$ and $(B_\ell)_{\ell \ge 0}$ are two sequences of nonnegative $n \times n$ matrices such that $A_\ell$ and $B_\ell$ have the same zero pattern for each $\ell$. Then for each $k \in \mathbb{N}$, the two matrices $\mathcal{A}_k = A_{k-1}A_{k-2}\cdots A_0$ and $\mathcal{B}_k = B_{k-1}B_{k-2}\cdots B_0$ share a zero pattern for any $k$.*

*Proof.* The proof is by induction. Suppose that $\mathcal{A}_k$ and $\mathcal{B}_k$ have the same zero pattern from some $k$. For any $i$,$j$, we know that $(\mathcal{A}_{k+1})_{ij} = 0$ if and only if, for each $l$, either $(A_k)_{il} = 0$ or $(\mathcal{A}_k)_{lj} = 0$. But this is the case if and only if $(B_k)_{il} = 0$ or $(\mathcal{B}_k)_{lj} = 0$ for each $l$, i.e., if and only if $(\mathcal{B}_{k+1})_{ij} = 0$. Since $A_0$ and $B_0$ have a common zero pattern, the result follows. $\square$

**Proposition 18.** *Let $k \in \mathbb{N}$ be given. For each $\ell \in \{0, 1, \dots, k-1\}$, let $A_\ell$ be a nonnegative $n \times n$ matrix, and let $\mathcal{A}_{\ell+1} = A_\ell A_{\ell-1} \cdots A_0$. Assume that both sets $\{(A_\ell)_{ij} : (A_\ell)_{ij} > 0,\ 0 \le \ell \le k-1\}$ and $\{(\mathcal{A}_k)_{ij} : (\mathcal{A}_k)_{ij} > 0\}$ are nonempty. Let*

$$\epsilon := \min_{i,j,\ell}\{(A_\ell)_{ij} : (A_\ell)_{ij} > 0,\ 0 \le \ell \le k-1\}.$$

*Then the following inequality holds true*:

$$\min_{i,j}\{(\mathcal{A}_k)_{ij} : (\mathcal{A}_k)_{ij} > 0\} \ge \epsilon^k.$$

*Proof.* We again prove by induction. The result is obviously true for $k = 1$. Now suppose that $k > 1$ and that for some $\ell < k$ we have $(\mathcal{A}_\ell)_{lm} \neq 0 \implies (\mathcal{A}_i)_{lm} \ge \epsilon^i$. We write down the $lm$-entry of the matrix $(\mathcal{A}_{\ell+1})$:

$$(\mathcal{A}_{\ell+1})_{lm} = \sum_{j=1}^{n} (A_{\ell+1})_{lj} (\mathcal{A}_\ell)_{jm}.$$

If $(\mathcal{A}_{\ell+1})_{lm}$ is positive, then there exists a $j$ such that both $(A_{\ell+1})_{lj}$ and $(\mathcal{A}_\ell)_{jm}$ are positive. Since $(A_{\ell+1})_{lj} \ge \epsilon$ and $(\mathcal{A}_\ell)_{jm} \ge \epsilon^\ell$, we have $(\mathcal{A}_{\ell+1})_{lm} \ge \epsilon^{\ell+1}$. That is, $(\mathcal{A}_{\ell+1})_{lm} \neq 0 \implies (\mathcal{A}_{\ell+1})_{lm} \ge \epsilon^{\ell+1}$. The induction process is complete. ☐

The following result is due to Tollisen and Lengyel [2008].

**Proposition 19.** *Let $A$ be an $n \times n$ circulant matrix with first row $(c_0, c_1, \dots, c_{n-1})$. Let $L = \{i : c_i > 0\}$, $u = \min L$, $L' = \{i - u : c_i > 0\}$, and $g = \gcd(L')$. Then*

$$(A^k)_{ij} \approx \begin{cases} \frac{1}{n}\gcd(n, g) & \text{if } j - i \equiv ku \pmod{\gcd(n, g)}, \\ 0 & \text{otherwise} \end{cases}$$

*as $k \to \infty$.*

The rest of this section is devoted to statements and proofs of the main result.

**Proposition 20.** *Let $A$ be a stochastic circulant matrix such that the sequence $A^k$ converges to a rank-one matrix $L$. Let $(A_\ell)_{\ell \ge 0}$ be a sequence of stochastic matrices having the same zero pattern as $A$. Moreover, assume that there exists an $\epsilon > 0$ such that*

$$\min_{i,j,\ell}\{(A_\ell)_{ij} : (A_\ell)_{ij} > 0\} \ge \epsilon.$$

*Then the sequence of matrices $A_k A_{k-1} \cdots A_1 A_0$ converges to a rank-one matrix $L'$ with identical rows.*

*Proof.* A result from [Kra and Simanca 2012] asserts that the product of circulant matrices is circulant. Hence $A^k$ is a sequence of stochastic circulant matrices, and so is their limit $L$. Proposition 19 assures us that each entry of $L$ is either zero or $\gcd(n, g)/n$. Suppose that for some $i$ and $j$, we have $(L)_{ij} = 0$. Since $L$ is stochastic, we must have $(L)_{ij'} > 0$ for some $j' \neq j$. Using arithmetic modulo $n$ to denote row and column indices, we can identify an $\ell$ such that $(L)_{i\ell} = 0$ and $(L)_{i,\ell+1} > 0$. Since $L$ is circulant, we have $(L)_{(i+1),(\ell+1)} = (L)_{i\ell} = 0$. Therefore, the $i$-th and the $(i+1)$-th rows of $L$ must be linearly independent. This contradicts the fact that $L$ is rank-one. Therefore the entries $(L)_{ij}$ are either all zero or all equal to $\gcd(n, g)/n$. That $L$ is stochastic rules out the former. In fact, all entries $(L)_{ij}$ are $1/n$, which implies that $\gcd(n, g) = 1$. It follows that, for $k$ sufficiently large, $A^k > 0$.

Define $\mathcal{B}_\ell = A_{\ell k-1} A_{\ell k-2} \cdots A_{(\ell-1)k}$ for $\ell \geq 1$. Then by Proposition 17, each $\mathcal{B}_\ell$ has the same zero pattern as $A^k$. That is, $\mathcal{B}_\ell > 0$ for all $\ell$. Furthermore, by Proposition 18 we know that $\mathcal{B}_\ell \geq \epsilon^k$. By Proposition 15, we have that

$$\tau_1(\mathcal{B}_\ell) \leq 1 - n\epsilon^k \quad \text{for } \ell \geq 0.$$

It follows that

$$\begin{aligned}\tau_1(A_\ell A_{\ell-1} \cdots A_0) &\leq \tau_1(\mathcal{B}_\ell \mathcal{B}_{\ell-1} \cdots \mathcal{B}_1) \\ &\leq \tau_1(\mathcal{B}_\ell)\tau_1(\mathcal{B}_{\ell-1}) \cdots \tau_1(\mathcal{B}_1) \leq (1 - n\epsilon^k)^\ell,\end{aligned}$$

which implies that the sequence $A_k A_{k-1} \cdots A_1 A_0$ converges to a rank-one matrix $L'$. Moreover, we can use the same Cauchy sequence argument as in the proof of Theorem 16 to show that the matrix $L'$ has identical rows. □

The following result is worth mentioning.

**Proposition 21.** *If $A$ is a stochastic circulant matrix, then $A^k$ converges to a rank-one matrix as $k \to \infty$ if and only if* $\gcd(n, g) = 1$.

*Proof.* On the one hand, as we observed in the previous proof, if $A$ is a stochastic circulant matrix such that $A^k$ converges to a rank-one matrix $L$ as $k \to \infty$, then $L$ must be strictly positive, and hence $\gcd(n, g) = 1$. On the other hand, if $A$ is a stochastic circulant matrix such that $\gcd(n, g) = 1$, then $j - i \equiv ku \pmod{\gcd(n, g)}$. Thus, $\lim_{k\to\infty}(A^k)_{ij} = 1/n$. That is, $A^k$ converges to the rank-one matrix whose entries are all $1/n$ as $k \to \infty$. □

We state the main result of this section in terms of convergent sequences of polygons.

**Theorem 22.** *Suppose that $(A_\ell)_{\ell\geq 0}$ is a sequence of stochastic, circulant-patterned, $n \times n$ matrices that all have a common zero pattern. Let* $\gcd(n, g) = 1$, *where* $u = \min\{i : a_i > 0\}$ *and* $g = \gcd\{i - u : a_i > 0\}$. *Here, $(a_0, a_1, \ldots, a_{n-1})$ is the first row of $A_0$. Assume that there exists an $\epsilon > 0$ such that*

$$\min_{i,j,k}\{(A_k)_{ij} : (A_k)_{ij} > 0\} \geq \epsilon.$$

*Then the sequence of matrices $A_k A_{k-1} \cdots A_0$ converges to a rank-one matrix $L$ that has identical rows. Hence, if $\Pi^{(k)} = A_k A_{k-1} \cdots A_1 \Pi^{(0)}$, then*

$$\lim_{k\to\infty} \Pi^{(k)} = L\Pi^{(0)}.$$

*That is, the sequence of polygons $(\Pi^{(k)})_{k\geq 0}$ converges to the point $L\Pi^{(0)}$.*

## Acknowledgment

The authors thank Alvin Moon, University of California, Davis, for calling their attention to G. Darboux's work, which predated Schoenberg's by more than 70 years.

## References

[Darboux 1878] G. Darboux, "Sur un problème de gèométrie élémentaire", *Bull. Sci. Math. Astron.* (2) **2**:1 (1878), 298–304.

[Ding et al. 2003] J. Ding, L. R. Hitt, and X.-M. Zhang, "Markov chains and dynamic geometry of polygons", *Linear Algebra Appl.* **367** (2003), 255–270. MR Zbl

[Horn and Johnson 1990] R. A. Horn and C. R. Johnson, *Matrix analysis*, corrected reprint of 1st ed., Cambridge University Press, 1990. Zbl

[Ipsen and Selee 2011] I. C. F. Ipsen and T. M. Selee, "Ergodicity coefficients defined by vector norms", *SIAM J. Matrix Anal. Appl.* **32**:1 (2011), 153–200. MR Zbl

[Kra and Simanca 2012] I. Kra and S. R. Simanca, "On circulant matrices", *Notices Amer. Math. Soc.* **59**:3 (2012), 368–377. MR Zbl

[Ouyang 2013] C. Ouyang, "A problem concerning the dynamic geometry of polygons", unpublished manuscript, 2013.

[Rosenman and Huston 1933] M. Rosenman and R. E. Huston, "3547", *The American Mathematical Monthly* **40**:3 (1933), 184–185. MR

[Schoenberg 1950] I. J. Schoenberg, "The finite Fourier series and elementary geometry", *Amer. Math. Monthly* **57** (1950), 390–404. MR Zbl

[Schoenberg 1982] I. J. Schoenberg, *Mathematical time exposures*, Mathematical Association of America, Washington, DC, 1982. MR Zbl

[Terras 1999] A. Terras, *Fourier analysis on finite groups and applications*, London Mathematical Society Student Texts **43**, Cambridge University Press, 1999. MR Zbl

[Tollisen and Lengyel 2008] G. P. Tollisen and T. Lengyel, "Intermediate and limiting behavior of powers of some circulant matrices", *Ars Combin.* **88** (2008), 229–255. MR Zbl

[Treatman and Wickham 2000] S. Treatman and C. Wickham, "Constructible approximations of regular polygons", *Amer. Math. Monthly* **107**:10 (2000), 911–922. MR Zbl

Received: 2015-03-04 Accepted: 2015-09-17

ehintikka@rice.edu *Department of Mathematics, MS 136, Rice University, P. O. Box 1892, Houston, TX 77005, United States*

xsun@missouristate.edu *Department of Mathematics, Missouri State University, Springfield, MO 65897, United States*

# involve

msp.org/involve


## INVOLVE YOUR STUDENTS IN RESEARCH

*Involve* showcases and encourages high-quality mathematical research involving students from all academic levels. The editorial board consists of mathematical scientists committed to nurturing student participation in research. Bridging the gap between the extremes of purely undergraduate research journals and mainstream research journals, *Involve* provides a venue to mathematicians wishing to encourage the creative involvement of students.




See inside back cover or msp.org/involve for submission instructions. The subscription price for 2016 is US $160/year for the electronic version, and $215/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

**mathematical sciences publishers**

**nonprofit scientific publishing**

http://msp.org/

# involve

2016 vol. 9 no. 5



1944-4176(2016)9:5;1-0