Can. J. Math., Vol. XXIX, No. 4, 1977, pp. 756–770

CYCLIC TRANSFORMATIONS OF POLYGONS AND THE GENERALIZED INVERSE

PHILIP J. DAVIS

1. Introduction; circulants. By a *polygon with n vertices* (an *n-gon*) we shall mean an ordered n-tuple of complex numbers $(z_1, z_2, \ldots, z_n)$. The numbers z_i will be called the *vertices* of the polygon. The z_i may or may not be distinct, and it shall not matter whether a "conventional" polygon can be drawn with the z_i as vertices. We shall operate in the spirit of the book of Bachmann and Schmidt [**2**] though not at the same level of algebraic generality inasmuch as these authors allow z_i to belong to a general field.

Let $(c_1, c_2, \ldots, c_n)$ be an ordered n-tuple of complex numbers, and let them generate the circulant matrix of order n:

$$(1.1) \qquad C = \begin{bmatrix} c_1 & c_2 & \ldots & c_n \\ c_n & c_1 & \ldots & c_{n-1} \\ & \cdot & & \cdot \\ & \cdot & & \cdot \\ & \cdot & & \cdot \\ c_2 & c_3 & \ldots & c_1 \end{bmatrix}$$

We shall often denote this matrix by

$$(1.2) \qquad C = \operatorname{circ}(c_1, c_2, \ldots, c_n).$$

Let the column vector $P = (z_1, \ldots, z_n)^T$ (T = transpose) symbolize the polygon P with vertices at z_i. We shall be considering cyclic transformations of P given by

$$(1.3) \qquad \tilde{P} = CP$$

where C is the circulant (1.2).

It is well known that all circulants of order n are simultaneously diagonalizable by the matrix F associated with the Finite Fourier Transform. Specifically, let

$$(1.4) \qquad \omega = \exp(2\pi i/n), \; i = \sqrt{-1},$$

Received August 11, 1976 and in revised form, November 23, 1976. Results shown in this paper were obtained, in part, in the course of research sponsored by the National Science Foundation under grant NSF-MCS-75-07468A01.

and set

$$(1.5)\qquad F^* = \frac{1}{\sqrt{n}}\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(n-1)} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & \omega^{n-1} & \omega^{n-2} & & \omega \end{bmatrix}.$$

The "Fourier matrix" F, depends only on n, is symmetric, unitary ($FF^* = F^*F = I$), and one has

$$(1.6)\qquad C = F^* \Lambda F$$

where

$$(1.7)\qquad \Lambda = \text{diag }(\lambda_1, \lambda_2, \dots, \lambda_n).$$

The symbol * designates the conjugate transpose. The eigenvalues λ_j of C are precisely

$$(1.8)\qquad \lambda_j = p(\omega^{j-1}), j = 1, 2, \dots, n$$

where $p(z)$ is the generating polynomial of C given by

$$(1.9)\qquad p(z) = c_1 + c_2 z + \dots + c_n z^{n-1}.$$

This may be seen from the spectral mapping theorem inasmuch as we may represent C in the form

$$(1.10)\qquad C = c_1 I + c_2 \Pi + c_3 \Pi^2 + \dots + c_n \Pi^{n-1}$$

where Π is the permutation matrix circ $(0, 1, 0, 0, \dots, 0)$ whose eigenvalues are $\omega^{j-1}, j = 1, 2, \dots, n$. Conversely, given a matrix C of the form (1.6), (1.7), by the fundamental theorem of polynomial interpolation, we can find a unique polynomial, $r(z)$, of degree $\leqq n - 1$: $r(z) = d_1 + d_2 z + \dots + d_n z^{n-1}$, such that $r(\omega^{j-1}) = \lambda_j$, $j = 1, 2, \dots, n$. Now form $D =$ circ $(d_1, d_2, \dots, d_n)$. It follows that $D = F^*\Lambda F = C$, so that C is a circulant.

From the representation (1.6), (1.7), it is easily shown that the sum, difference, scalar product, product, transpose, conjugate transpose, of circulants of order n is again a circulant of order n. Moreover, all circulants of the same order commute.

Since the rank of a diagonalizable matrix is equal to the number of its non-zero eigenvalues, and since from (1.6), $C^k = F^*\Lambda^k F$, it follows that

$$(1.11)\qquad \text{rank } C = \text{rank } C^k, \quad k = 1, 2, \dots$$

Note also that

$$(1.12)\qquad (\text{circ }(c_1, c_2, \dots, c_n))^T = \text{circ }(c_1, c_n, c_{n-1}, \dots, c_2).$$

If a circulant is non-singular, its inverse is also a circulant.

We turn next to generalized inverses for circulants. If M is a square matrix, and if there is a square matrix $M^{\#}$ of the same order satisfying

(a) $MM^{\#}M = M$
(b) $M^{\#}MM^{\#} = M^{\#}$
(c) $MM^{\#} = M^{\#}M$,

then $M^{\#}$ is called the *group inverse* of M. (See Drazin [**7**], Englefield [**8**], Erdelyi [**9**], and Ben-Israel and Greville [**3**], Chap. 4.) It is known that $M^{\#}$ exists if and only if rank M = rank M^2, and that if it exists, it is unique. Since circulants satisfy this rank equation (see 1.11), it follows that every circulant C has a unique group inverse $C^{\#}$.

For scalar λ set

$$(1.13)\quad \begin{cases} \lambda^{\div} = 1/\lambda & \text{for } \lambda \neq 0 \\ \lambda^{\div} = 0 & \text{for } \lambda = 0, \end{cases}$$

and for $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, set

$$(1.14)\quad \Lambda^{\div} = \operatorname{diag}(\lambda_1^{\div}, \lambda_2^{\div}, \ldots, \lambda_n^{\div}).$$

Then, the group inverse of the circulant $C = F^*\Lambda F$ is the circulant

$$(1.15)\quad C^{\#} = F^*\Lambda^{\div}F,$$

for it is easily verified that the right hand of (1.15) satisfies (a), (b), (c). Insofar as all circulants commute, it follows that condition (c) is redundant and we may say that the group inverse of the circulant C is the circulant $C^{\#}$ which satisfies (a) and (b).

Given an $m \times n$ matrix M, there is always a unique $n \times m$ matrix $M^{\div}$ which satisfies the four conditions

(a) $MM^{\div}M = M$
(b) $M^{\div}MM^{\div} = M^{\div}$
(d) $(MM^{\div})^* = MM^{\div}$
(e) $(M^{\div}M)^* = M^{\div}M$.

The matrix $M^{\div}$ is known as the *Moore-Penrose generalized inverse of* M ($M - P$ *inverse*). For a circulant $C = F^*\Lambda F$, it is also easily verified by direct substitution that the circulant $C^{\#} = F^*\Lambda^{\div}F$ satisfies (a), (b), (d), (e). Hence, for circulants, the group inverse and the $M - P$ inverse are identical:

$$(1.16)\quad C^{\#} = C^{\div}.$$

(In general, for $M^{\#}$ to equal $M^{\div}$ it is necessary and sufficient for M and $M^{\div}$ to commute.)

In the present paper, we stress the $M - P$ inverse because of its least square properties, shortly to be noted, which play a role in Sections 3, 5 and 6. In addition, several computer languages, such as APL, return $M^{\div}$ when the inverse of M is called for, so that in experimental numerical work one tends to think along $M - P$ lines.

The system of linear equations

(1.17) $MX = B$,

where X and B are columns, has a solution if and only if $MM^{\div}B = B$. If this is the case, then the general solution of (1.17) is given by

(1.18) $X = M^{\div}B + (I - M^{\div}M)Y$,

where Y is an arbitrary $n \times 1$ matrix. (This is true if the generalized inverse satisfies only (a).)

For vectors $Z = (z_1, z_2, \ldots, z_n)$, introduce the norm

$$||Z||^2 = \sum_{k=1}^{n} |z_k|^2.$$

Given a system of linear equations $MX = B$, if the columns of M are independent then the least square problem

(1.19) $||MX - B|| = \text{minimum}$

has the unique solution

(1.20) $X = M^{\div}B$.

If they are not, the problem (1.19) has many solutions of which (1.20) is the unique solution of minimum norm.

The spectral decomposition of a circulant should be noted. Let $C = F^*\Lambda F$ where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Introduce the diagonal matrices

(1.21) $\Lambda_k = \text{diag}(0, 0, \ldots, 0, 1, 0, \ldots, 0)$

$k = 1, 2, \ldots, n$, where the 1 occurs in the kth position. Now $\Lambda = \sum_{k=1}^{n} \lambda_k \Lambda_k$, so that $C = \sum_{k=1}^{n} \lambda_k F^*\Lambda_k F$. If we set

(1.22) $B_k = F^*\Lambda_k F, \quad k = 1, 2, \ldots, n,$

then we can write

(1.23) $C = \sum_{k=1}^{n} \lambda_k B_k.$

Note that $B_j B_k = F^*\Lambda_j F F^*\Lambda_k F = F^*OF = O$ if $j \neq k$, while $B_k^2 = F^*\Lambda_k\Lambda_k F^* = F^*\Lambda_k F = B_k$. Also, $B_k^* = F^*\Lambda_k^* F = F^*\Lambda_k F = B_k$. In the special case where $\Lambda = \text{diag}(1, 1, \ldots, 1) = I$, $C = F^*F = I = \sum_{k=1}^{n} B_k$, so that the B_k are projections and constitute a resolution of unity. For $p = 0, 1, \ldots$ it follows that

(1.24) $C^p = \sum_{k=1}^{n} \lambda_k^p B_k.$

https://doi.org/10.4153/CJM-1977-079-8 Published online by Cambridge University Press

If C is non-singular, (1.24) persists for all integer p: $-\infty < p < \infty$, while if C is singular one has

(1.25) $$C^{\div} = \sum_{k=1}^{n} \lambda_k^{\div} B_k.$$

In particular,

(1.26) $$B_k^{\div} = B_k.$$

2. Circulants of rank $n - 1$. Insofar as a circulant is diagonalizable, a circulant of rank $n - 1$ has precisely one zero eigenvalue. If $C = F^*\Lambda F$, then C has rank $n - 1$ if and only if for some integer j, $1 \leqq j \leqq n$,

(2.1) $$\Lambda = \operatorname{diag}(u_1, \ldots, u_{j-1}, 0, u_{j+1}, \ldots, u_n),$$

with $u_i \neq 0$, $i \neq j$. Now,

(2.2) $$\Lambda^{\div} = \operatorname{diag}(u_1^{-1}, \ldots, u_{j-1}^{-1}, 0, u_{j+1}^{-1}, \ldots, u_n^{-1})$$

and $C^{\div} = F^*\Lambda^{\div}F$. Thus, we have,

(2.3) $$CC^{\div} = C^{\div}C = F^* \operatorname{diag}(1, 1, \ldots, 1, 0, 1, \ldots, 1)F$$

where the 0 occurs in the jth position. From this it follows that

(2.4) $$CC^{\div} = F^*(I - \Lambda_j)F = I - F^*\Lambda_j F = I - B_j, \quad \text{or}$$

(2.5) $$B_j = I - CC^{\div}.$$

Let $\omega = \exp(2\pi i/n)$, $i = \sqrt{-1}$, and set

(2.6) $$K_r = (1/n) \operatorname{circ}(1, \omega^r, \omega^{2r}, \ldots, \omega^{(n-1)r}).$$

The generating polynomial for K_r is $p(z) = (1/n)(1 + \omega^r z + \omega^{2r}z^2 + \ldots + \omega^{(n-1)r}z^{n-1}) = ((\omega^r z)^n - 1)/n(\omega^r z - 1)$. The eigenvalues of K_r are $p(\omega^{j-1})$, $j = 1, 2, \ldots, n$. Now for $j - 1 \neq n - r$, $p(\omega^{j-1}) = 0$, while $p(\omega^{n-r+1}) = 1$. Thus, if

(2.7) $$r = n - j + 1,$$

then $K_r = F^* \operatorname{diag}(0, 0, \ldots, 0, 1, 0, \ldots, 0)F$, the 1 occurring in the jth position. This means that

(2.8) $$K_r = F^*\Lambda_j F = B_j.$$

Thus,

(2.9) $$I - CC^{\div} = K_r.$$

Notice that

(2.10) $$K_r^2 = B_j^2 = B_j = K_r; \quad K_r K_s = O, \ r \neq s.$$

From (2.5), (2.8), (a), (b),

$$\begin{cases} CK_r = CB_j = C(I - CC^{\div}) = C - CCC^{\div} = O \\ C^{\div}K_r = C^{\div}B_j = C^{\div}(I - CC^{\div}) = C^{\div} - C^{\div}CC^{\div} = O. \end{cases} \tag{2.11}$$

Several more identities will be of use. Again, let $K_r = (1/n)$ circ $(1, \omega^r, \omega^{2r}, \ldots, \omega^{(n-1)r})$. Let Y be an arbitrary circulant so that one can write $Y = F^* \operatorname{diag}(\eta_1, \eta_2, \ldots, \eta_n)F$ for appropriate η_i. Now, $K_rY = (F^*\Lambda_jF)(F^* \operatorname{diag}(\eta_1, \ldots, \eta_n)F) = F^* \operatorname{diag}(0, \ldots, 0, \eta_j, 0, \ldots, 0)F = \eta_jF^*\Lambda_jF = \eta_jK_r$. Thus,

$$K_rY = \eta_jK_r. \tag{2.12}$$

In particular, if Y is merely a column vector $Y = (y_0, y_1, \ldots, y_{n-1})^T$, then

$$K_rY = \eta_j \operatorname{fc}(K_r) \tag{2.13}$$

where the notation fc (K_r) designates the first column of K_r. One also has

$$K_rY = \sigma(1, \omega^{(n-1)r}, \omega^{(n-2)r}, \ldots, \omega^r)^T \tag{2.14}$$

where

$$\sigma = y_0 + y_1\omega^r + \ldots + y_{n-1}\omega^{(n-1)r}. \tag{2.15}$$

Let Y be further specialized to $Y = \operatorname{fc}(K_r)$. Then $Y = (1/n)(1, \omega^{(n-1)r}, \omega^{(n-2)r}, \ldots, \omega^r)^T$. Therefore from (2.15), $\sigma = 1$, and from (2.14),

$$K_r \operatorname{fc}(K_r) = \operatorname{fc}(K_r). \tag{2.16}$$

3. n-gons and K_r-grams. In what follows we shall assume that C is a circulant of order n and of rank $n - 1$. Then, as we have seen, $CC^{\div} = I - K_r$ for a unique integer r, $1 \leqq r \leqq n$.

Let $P = (z_1, z_2, \ldots, z_n)^T$ be an n-gon. We shall say that P is a *K_r-gram* if and only if

$$z_1 + \omega^r z_2 + \omega^{2r}z_3 + \ldots + \omega^{(n-1)r}z_n = 0. \tag{3.1}$$

Insofar as all the rows of K_r are identical to the first row $1, \omega^r, \ldots, \omega^{(n-1)r}$ multiplied by some ω^l, it follows that *P is a K_r-gram if and only if*

$$K_rP = 0. \tag{3.2}$$

Each circulant C of rank $n - 1$ determines a class of K_r-grams.

THEOREM. *Let P be an n-gon. Then, there exists an n-gon $\hat{P}$ such that $C\hat{P} = P$ if and only if P is a K_r-gram.*

Proof. The system of equations $C\hat{P} = P$ has a solution if and only if $P = CC^{\div}P$. This is equivalent to $P = (I - K_r)P = P - K_rP$ or $K_rP = 0$ (by (2.9)).

COROLLARY. *Let P be a K_r-gram. Then the general solution to $C\hat{P} = P$ is given by*

$$(3.3) \qquad \hat{P} = C^{\div}P + \tau \text{ fc } (K_r)$$

for an arbitrary constant τ.

Proof. If P is a K_r-gram, then the general solution to $C\hat{P} = P$ is given by $\hat{P} = C^{\div}P + (I - C^{\div}C)Y = C^{\div}P + K_rY$ for an arbitrary column vector Y. From (2.13), $K_rY = \eta_j \text{ fc } (K_r)$ and the statement follows.

COROLLARY. *P is a K_r-gram if and only if there is an n-gon Q such that $P = CQ$.*

Proof. Let $P = CQ$. Then $K_rP = K_rCQ$. Since $K_rC = O$, it follows that $K_rP = O$ so that P is a K_r-gram. Conversely, let P be a K_r-gram. Now take for Q any $\hat{P}$ whose existence is guaranteed by the previous corollary.

COROLLARY. *Given an n-gon P which is a K_r-gram. Then, given an arbitrary complex number $\hat{z}_1$, we can find a unique n-gon $\hat{P} = (\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_n)^T$, with $\hat{z}_1$ as its first vertex and such that $C\hat{P} = P$.*

Proof. Since the general solution of $C\hat{P} = P$ is $P = C^{\div}P + \tau \text{ fc } (K_r)$, given $\hat{z}_1$, we may solve uniquely for an appropriate τ since the first component of fc (K_r) is 1 $(\neq 0)$.

THEOREM. *Let P be an n-gon which is a K_r-gram. Then, there is a unique n-gon Q which is a K_r-gram and such that $CQ = P$. It is given by $Q = C^{\div}P$.*

Proof. (a) Since P is a K_r-gram, it has the form $P = CR$ for some R. Hence $Q = C^{\div}P = C^{\div}CR = C(C^{\div}R)$. Hence Q is a K_r-gram. (b) Q is a solution of $CQ = P$, as we can see by selecting $\tau = 0$ in the above. (c) All solutions are of the form $\hat{P} = C^{\div}P + \tau \text{ fc } (K_r)$. Now $\hat{P}$ is a K_r-gram if and only if $K_r\hat{P} = 0$. That is, if and only if $K_rC^{\div}P + \tau K_r \text{ fc } (K_r) = 0$. Now $K_rC^{\div} = O$. But K_r fc $(K_r) = K_r$. Therefore $\tau = 0$.

THEOREM. *Let P be a K_r-gram. Amongst the infinitely many n-gons R for which $CR = P$, there is a unique one of minimum norm $||R||$. It is given by $R = C^{\div}P$. Hence it coincides with the unique K_r-gram Q such that $CQ = P$.*

Proof. Use the last theorem and the least squares characterization of the $M - P$ inverse.

Suppose now that P is a general n-gon and we wish to approximate it by a K_r-gram R such that $||P - R|| =$ minimum. Every K_r-gram can be written as $R = CQ$ for some n-gon Q, so that our problem is: given P, find a Q such that $||P - CQ|| =$ minimum. This problem has a solution, and the solution is unique if and only if the columns of C are linearly independent. This is not the case (the rank of C being $n - 1$), hence $Q = C^{\div}P$ is the solution with minimum $||Q||$. Thus, $R = CQ = CC^{+}P$ is the best approximation of the n-gon P by a K_r-gram with minimum $||Q||$. We phrase this as follows:

THEOREM. *Given a general n-gon $P = (z_1, \ldots, z_n)^T$. The unique K_r-gram $R = CQ$ for which $||P - R|| =$ minimum and $||Q|| =$ minimum is given by*

$$(3.4) \quad R = CC^{+}P = (I - K_r)P = P - K_rP$$
$$= P - \sigma(1, \omega^{(n-1)r}, \omega^{(n-2)r}, \ldots, \omega^r)^T$$

where $\sigma = z_1 + z_2\omega^r + \ldots + z_n\omega^{(n-1)r}$. Alternatively, this can be written as

$$(3.5) \quad R = P - \eta_j \operatorname{fc}(K_r)$$

where η_j is determined from

$$\operatorname{circ}(z_1, z_2, \ldots, z_n) = F^* \operatorname{diag}(\eta_1, \eta_2, \ldots, \eta_n)F.$$

Proof. As before, $R = CC^{+}P = (I - K_r)P = P - K_rP$. By (2.13) $K_rP = \eta_j \operatorname{fc}(K_r)$. Notice that R is a K_r-gram because $K_rR = K_r(P - \eta_j \operatorname{fc}(K_r)) = K_rP - \eta_jK_r \operatorname{fc}(K_r)$. Since by (2.16), $K_r \operatorname{fc}(K_r) = \operatorname{fc}(K_r)$, $K_rR = 0$.

Notice also that if P is already a K_r-gram, $\sigma = z_1 + z_2\omega^r + \ldots + z_n\omega^{(n-1)r} = 0$. In this case, from (3.4), $R = P$; so as expected, P is its own best approximation.

Generally, of course, the operation $R(P) = CC^{+}P$ is a projection onto the row or column space of C.

4. Further restrictions; geometric interpretations. An interesting class of cyclic transformations comes about from circ $(s, t, 0, 0, \ldots, 0)$, of order n, where one assumes that $s + t = 1$, $st \neq 0$, and that the rank is $n - 1$. Write

$$(4.1) \quad C_s = \operatorname{circ}(s, 1 - s, 0, 0, \ldots, 0).$$

The generating polynomial is $p(z) = s + (1 - s)z$, so that the eigenvalues of C_s are $p(\omega^k) = s + (1 - s)\omega^k$, $k = 0, 1, \ldots, n - 1$. Suppose that for a fixed j, $0 \leqq j \leqq n - 1$, $s + (1 - s)\omega^j = 0$. Thus, there will be a zero eigenvalue if and only if $s = \omega^j/(\omega^j - 1)$, $t = 1/(1 - \omega^j)$. For such s, C_s can have no more than one zero eigenvalue since $s + (1 - s)\omega^k = s + (1 - s)\omega^j = 0$ implies that $\omega^k = \omega^j$, or $k = j$. Thus, we have

THEOREM. *The circulant C_s has rank $n - 1$ if and only if for some integer j, $0 \leqq j \leqq n - 1$,*

$$(4.2) \quad s = \omega^j/(\omega^j - 1), \quad 1 - s = 1/(1 - \omega^j).$$

In this case,

$$(4.3) \quad C_sC_s^{+} = I - K_{n-j}.$$

If s is real, then C_s has rank $n - 1$ if and only if n is even and $s = t = 1/2$.

Proof. The $j + 1$st eigenvalue of C_s is zero. Hence (4.2) follows by (2.7), (2.9). If s is real, so is $1 - s$ and hence $1 - \omega^j$. Therefore ω^j = real. Since $j = 0$ is impossible ($s = \infty$), $\omega^j = -1$. This can happen if and only if n is even. From (4.2), $s = t = 1/2$.

If s is real, the transformation induced by C_s is interesting visually because the vertices of $\tilde{P} = C_sP$ lie on the sides (possibly extended) of P. Moreover, if s and t are limited by

(4.4) $\quad s + t = 1, s > 0, t > 0$

i.e., a convex combination, then $\tilde{P}$ is obtained from P in a simple manner: the vertices of $\tilde{P}$ divide the sides of P internally into the ratio $s : 1 - s$. Figure 1 is drawn for $s = 2/3$, $t = 1/3$.

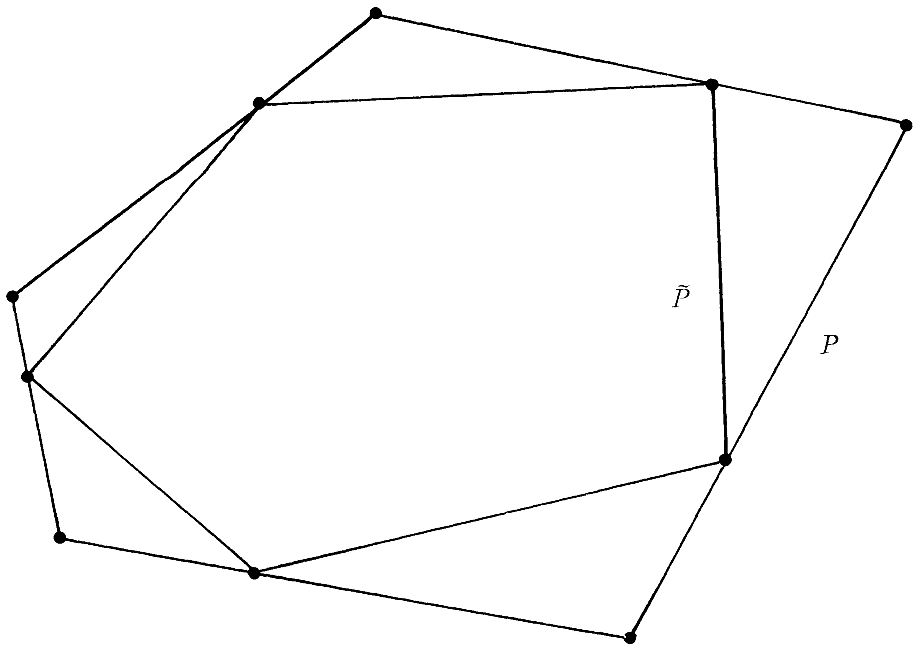

FIGURE 1

(For a discussion of the convergence of the iteration of this process, see I. J. Schoenberg [7].)

If s and t are complex, we shall point out the geometric interpretation subsequently.

As seen, if n = even and s is real, then C_s is singular if and only if $s = t = 1/2$. In all other real cases, the circulant C_s is non-singular and hence, given an arbitrary n-gon P, it will have a unique pre-image $\hat{P}$ under C_s: $C_s\hat{P} = P$.

https://doi.org/10.4153/CJM-1977-079-8 Published online by Cambridge University Press

Example. Let $n = 4$, $s = t = 1/2$. If Q is any quadrilateral, then $C_{1/2}Q$ is obtained from Q by joining successively the midpoints of the sides of Q. It is therefore a parallelogram. Hence, if one starts a quadrilateral Q, which is not a parallelogram, it can have no pre-image under $C_{1/2}$.

Since in such a case the system of equations can be "solved" by the application of a generalized inverse, we seek a geometric interpretation of this process.

5. The parallelogram and the Moore-Penrose inverse. Select $n =$ even, $s = t = 1/2$. Then, $C_s = \text{circ}\ (\frac{1}{2}, \frac{1}{2}, 0, \dots, 0)$. For simplicity designate $C_{1/2}$ by D:

(5.1) $$D = \text{circ}\ (\tfrac{1}{2}, \tfrac{1}{2}, 0, \dots, 0).$$

This corresponds to $j = n/2$ in (4.2). Hence by (4.3),

(5.2) $$DD^{\div} = I - K_{n/2}$$

where by (2.6)

(5.3) $$K_{n/2} = (1/n)\ \text{circ}\ (1, -1, 1, -1, \dots, 1, -1).$$

For simplicity we write $K_{n/2} = K$.

It is of some interest to have the explicit expression for $D^{\div}$.

THEOREM. *Let* $D = \text{circ}\ (\frac{1}{2}, \frac{1}{2}, 0, 0, \dots, 0)$ *be of order* n, *where* n *is even. Let*

(5.4) $$E = \text{circ}\,\frac{(-1)^{(n/2)-1}}{n}\ ((-1)^{(n/2)-1}(n-1), \dots, 5, -3, 1, 1, -3, 5, \dots, (-1)^{(n/2)-1}(n-1))$$

Then $E = D^{\div}$.

As particular instances note:

$$n = 4\colon D^{\div} = \text{circ}\ \tfrac{1}{4}(3, -1, -1, 3)$$
$$n = 6\colon D^{\div} = \text{circ}\ \tfrac{1}{6}(5, -3, 1, 1, -3, 5).$$

Proof. (1). A simple computation shows that

$$DE = \text{circ}\ (1/n)(n-1, 1, -1, 1, -1, \dots, -1, 1) = I - K.$$

Hence, $DED = (I - K)D = D - KD = D$, since by (2.11) (or by a direct computation) $KD = O$.

(2) On the other hand, $EDE = DEE = (I - K)E = E - KE$. An equally simple computation shows that $KE = O$. Hence $EDE = E$.

Thus by the remarks following (1.15), $E = D^{\div}$.

From (3.1) or (3.2), in the case under study, a K-gram is an n-gon whose vertices $z_1, \dots, z_n$ satisfy

(5.5) $$z_1 - z_2 + z_3 - z_4 + \dots + z_{n-1} - z_n = 0.$$

https://doi.org/10.4153/CJM-1977-079-8 Published online by Cambridge University Press

It is easily verified that for $n = 4$ the condition

$$(5.5') \quad z_1 - z_2 + z_3 - z_4 = 0$$

holds if and only if z_1, z_2, z_3, z_4 (in that order) form a conventional parallelogram. Thus, an n-gon which satisfies (5.5) is a "generalized" parallelogram. The sequence of theorems of Section 3 can now be given specific content in terms of parallelograms or generalized parallelograms. We shall write it up in terms of parallelograms.

THEOREM. *Let P be a quadrilateral. Then, there exists a quadrilateral $\hat{P}$ such that $D\hat{P} = P$ (the midpoint property) if and only if P is a parallelogram.*

COROLLARY. *Let P be a parallelogram. Then the general solution to $D\hat{P} = P$ is given by*

$$(5.6) \quad \hat{P} = D^{\div}P + \tau(1, -1, 1, -1)^T$$

for an arbitrary constant τ.

COROLLARY. *P is a parallelogram if and only if there is a quadrilateral Q such that $P = DQ$.*

COROLLARY. *Let P be a parallelogram. Then, given an arbitrary number $\hat{z}_1$ we can find a unique quadrilateral $\hat{P}$ with $\hat{z}_1$ as its first vertex such that $D\hat{P} = P$.*

THEOREM. *Let P be a parallelogram. Then there is a unique parallelogram Q such that $DQ = P$. It is given by $Q = D^{\div}P$.*

Notice what this is saying. DQ is the parallelogram formed from the midpoints of the sides of Q. Given a parallelogram P, we can find infinitely many quadrilaterals Q such that $DQ = P$. The first vertex may be chosen arbitrarily and this fixes all other vertices uniquely. But there is a unique *parallelogram* Q such that $DQ = P$. It can be found from $Q = D^{\div}P$ (see Figure 2).

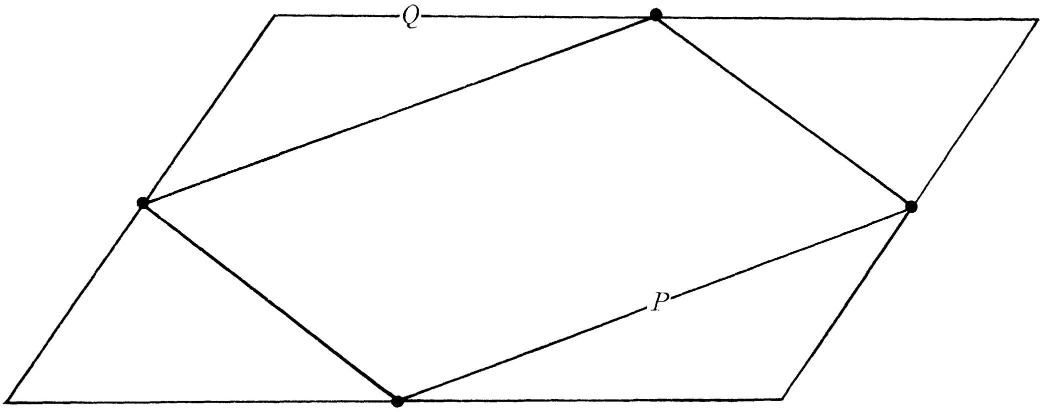

FIGURE 2

https://doi.org/10.4153/CJM-1977-079-8 Published online by Cambridge University Press

THEOREM. *Let P be a parallelogram. Amongst the infinitely many quadrilaterals R for which $DR = P$, there is a unique one of minimum norm $||R||$. It is given by $R = D^{+}P$. Hence it coincides with the unique parallelogram Q such that $DQ = P$.*

THEOREM. *Let P be a general quadrilateral. The unique parallelogram $R = DQ$ for which $||P - R||$ = minimum and $||Q||$ = minimum is given by $R = (I - K)P$.*

6. Napoleon's theorem and the Moore-Penrose inverse. The theorem in question is: if equilateral triangles are erected outwardly (or inwardly) upon the sides of *any* triangle, then their centers will form an equilateral triangle. This theorem, strangely, has been attributed to Napoleon Bonaparte. The reader is referred to Coxeter [**5**], Coxeter and Greitzer [**6**] for further references and facts. The triangles resulting are called the *outer* (or *inner*) Napoleon triangles corresponding to the given triangle. See Figure 3 where the outer case is depicted.

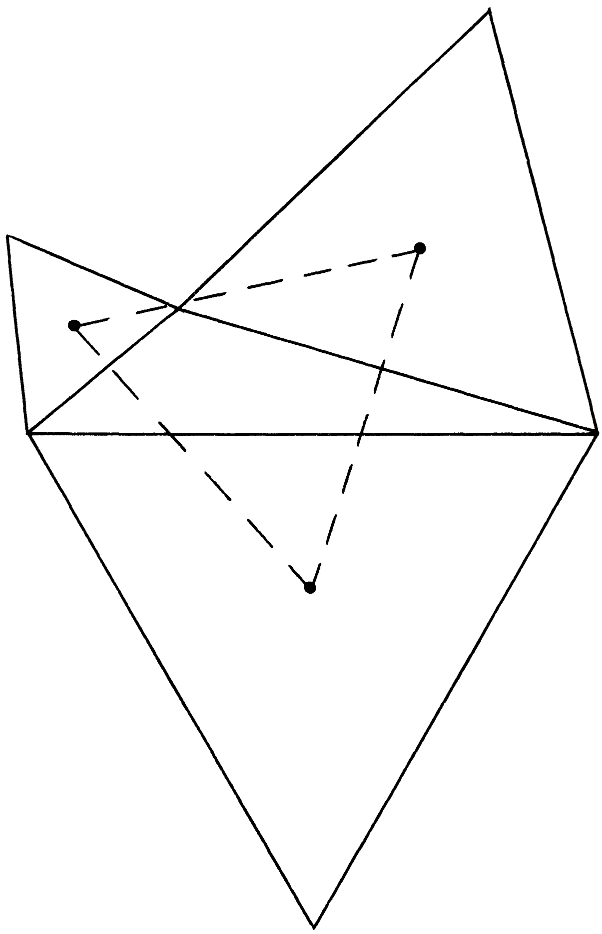

FIGURE 3

https://doi.org/10.4153/CJM-1977-079-8 Published online by Cambridge University Press

In the theorem of Section 4, select $n = 3$, and $\omega = \exp(2\pi i/3)$, so that $\omega^3 = 1$. Select $j = 1$, so that $s = \omega/(\omega - 1)$, $1 - s = 1/(1 - \omega)$. In view of $1 + \omega + \omega^2 = 0$, this simplifies to $s = \frac{1}{3}(1 - \omega)$, $1 - s = \frac{1}{3}(1 - \omega^2)$. On the other hand, the selection $j = 2$ leads to $s = \omega^2/(\omega^2 - 1) = \frac{1}{3}(1 - \omega^2)$, $1 - s = 1/(1 - \omega^2) = \frac{1}{3}(1 - \omega)$. The corresponding circulants C_s we shall designate by N (in honor of Napoleon):

$$\text{(6.1)} \qquad \begin{aligned} N_I &= \text{circ } \tfrac{1}{3}(1 - \omega, 1 - \omega^2, 0), \quad j = 1 \\ N_O &= \text{circ } \tfrac{1}{3}(1 - \omega^2, 1 - \omega, 0), \quad j = 2 \end{aligned}$$

the subscripts I, O standing for "inner" and "outer". For brevity we exhibit only the outer case, writing

$$\text{(6.1}') \qquad N = \text{circ } \tfrac{1}{3}(1 - \omega^2, 1 - \omega, 0).$$

We have

$$\text{(6.2)} \qquad \begin{aligned} K_0 &= \text{circ } \tfrac{1}{3}(1, 1, 1) \\ K_1 &= \text{circ } \tfrac{1}{3}(1, \omega, \omega^2), \quad K_0 + K_1 + K_2 = I \\ K_2 &= \text{circ } \tfrac{1}{3}(1, \omega^2, \omega). \end{aligned}$$

From (4.3) with $n = 3, j = 2$,

$$\text{(6.3)} \qquad NN^{\div} = I - K_1$$

THEOREM. $N^{\div} = K_0 - \omega K_2$.

Proof. Let $E = K_0 - \omega K_2$. Then from (6.1′), $N = K_0 - \omega^2 K_2$. Hence, $NE = (K_0 - \omega^2 K_2)(K_0 - \omega K_2) = K_0^2 + \omega^3 K_2^2 = K_0 + K_2 = I - K_1$ (cf. (2.10)). Therefore $NEN = (I - K_1)(K_0 - \omega^2 K_2) = K_0 - \omega^2 K_2 = N$. Similarly, $ENE = (I - K_1)(K_0 - \omega K_2) = K_0 - \omega K_2 = E$. Thus, by the remarks following (1.15), $E = N^{\div}$.

Let z_1, z_2, z_3 (in this order) be the vertices of an equilateral triangle described counter-clockwise. Since the direct conformal transformation $t(z) = az + b$ is a rotation and stretching followed by a translation, z_1, z_2, z_3 must be the images under some $t(z)$ of $1, \omega, \omega^2$. It follows from this that

$$\text{(6.4)} \qquad z_1 + \omega z_2 + \omega^2 z_3 = 0$$

is necessary and sufficient that z_1, z_2, z_3 be counterclockwise equilateral while

$$\text{(6.4}') \qquad z_1 + \omega^2 z_2 + \omega z_3 = 0$$

is necessary and sufficient that z_1, z_2, z_3 be clockwise equilateral. Now (6.4) can be written as $K_1(z_1, z_2, z_3)^T = 0$ while (6.4′) is $K_2(z_1, z_2, z_3)^T = 0$. Therefore a counter-clockwise equilateral triangle is a K_1-gram, while a clockwise equilateral triangle is a K_2-gram.

Let now (z_1, z_2, z_3) be the vertices of an arbitrary triangle. On the sides of this triangle erect equilateral triangles outwardly. Let their vertices be z_1', z_2', z_3'. From (6.4),

$$z_1' = -\omega^2 z_1 - \omega z_2, \quad z_2' = -\omega^2 z_2 - \omega z_3, \quad z_3' = -\omega z_1 - \omega^2 z_3.$$

The centers of these equilateral triangles are therefore

$$(6.5) \quad z_1'' = \tfrac{1}{3}(1 - \omega^2)z_1 + \tfrac{1}{3}(1 - \omega)z_2, \quad z_2'' = \tfrac{1}{3}(1 - \omega^2)z_2 + \tfrac{1}{3}(1 - \omega)z_3,$$
$$z_3'' = \tfrac{1}{3}(1 - \omega^2)z_3 + \tfrac{1}{3}(1 - \omega)z_1.$$

This may be written as

$$(6.5') \quad (z_1'', z_2'', z_3'')^T = N(z_1, z_2, z_3)^T,$$

providing us with a geometric interpretation of the transformation induced by Napoleon's matrix.

The sequence of theorems of Section 3 can now be given specific content in terms of the Napoleon operator. In what follows all figures are taken counterclockwise.

THEOREM. *Let T be a triangle. Then there exists a triangle $\hat{T}$ such that $N\hat{T} = T$ if and only if T is equilateral.*

(The "only if" part is Napoleon's Theorem.)

COROLLARY. *Let T be equilateral. Then, the general solution to $N\hat{T} = T$ is given by*

$$(6.6) \quad \hat{T} = N^{\div}T + \tau(1, \omega^2, \omega)^T$$

for an arbitrary constant τ.

COROLLARY. *T is equilateral if and only if $T = NQ$ for some triangle Q.*

COROLLARY. *Given an equilateral triangle T. Given also an arbitrary complex number $\hat{z}_1$. There is a unique triangle $\hat{T}$ with $\hat{z}_1$ as its first vertex such that $N\hat{T} = T$.*

THEOREM. *Let T be an equilateral triangle. Then, there is a unique equilateral triangle Q such that $NQ = T$. It is given by $Q = N^{\div}T$.*

THEOREM. *Let T be equilateral. Let R be any triangle with $NR = T$. The unique such R of minimum norm $||R||$ is the equilateral triangle $R = N^{\div}T$. It is identical to the unique equilateral triangle Q for which $NQ = T$.* (See Figure 4).

Finally, suppose we are given an arbitrary triangle T and we wish to approximate it optimally by an equilateral triangle. Here is the story.

THEOREM. *Let T be arbitrary, then the equilateral triangle NR for which $||T - NR|| =$ minimum and such that $||R|| =$ minimum is given by $R = N^{\div}T$ and $NR = NN^{\div}T = (I - K_1)T$.*

Acknowledgement. The author wishes to thank the referee for a number of perceptive observations.

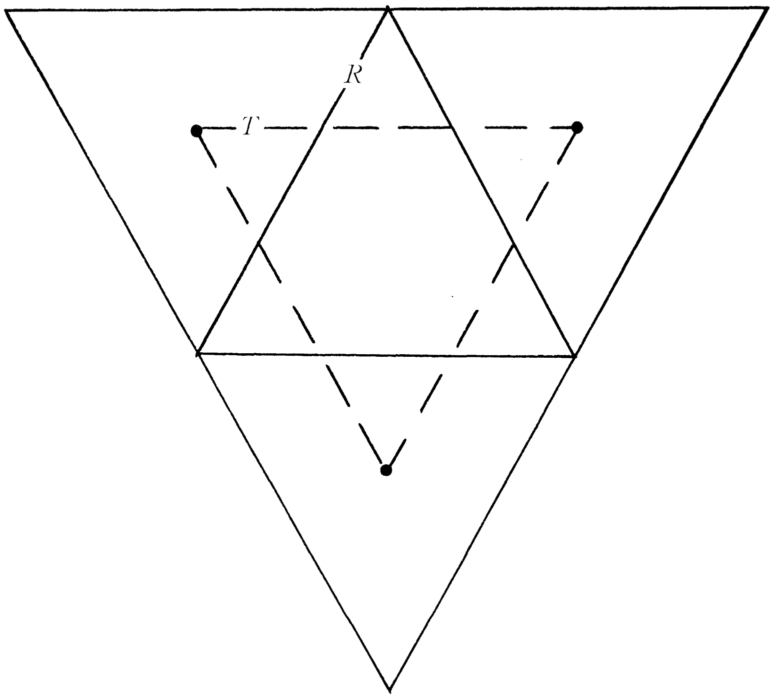

FIGURE 4

REFERENCES

1. J. H. Ahlberg, *Block circulants of level K*, Division of Applied Mathematics, Brown University, Prov. R.I., June, 1976.
2. F. Bachmann and E. Schmidt, *n-gons*, Translated by C. W. L. Garner, Mathematical Exposition No. 18 (Univ. of Toronto Press, Toronto, 1975).
3. A. Ben-Israel and T. N. E. Greville, *Generalized inverses* (Academic Press, N.Y., 1974).
4. G. Choquet, *Geometry in a modern setting* (London, 1969).
5. H. S. M. Coxeter, *Introduction to geometry* (Wiley, N. Y., 1969).
6. H. S. M. Coxeter and S. L. Greitzer, *Geometry revisited* (Mathematical Association of America, 1975).
7. M. P. Drazin, *Pseudo inverses in associative rings and semigroups*, Amer. Math. Monthly *65* (1958), 506–514.
8. M. J. Englefield, *The commuting inverses of a square matrix*, Proc. Camb. Phil. Soc. *62* (1966), 667–671.
9. I. Erdelyi, *On the matrix equation $Ax = \lambda Bx$*, J. Math. Anal. Appl. *17* (1967), 119–132.
10. I. J. Schoenberg, *The finite Fourier series and elementary geometry*, Amer. Math. Monthly *57* (1950), 390–404.
11. G. E. Trapp, *Inverses of circulant matrices and block circulant matrices*, Kyungpook Math. Jour. *13* (1973), 11–20.

Brown University,
Providence, Rhode Island