Reprinted from the MATHEMATICS MAGAZINE
Vol. 48, No. 2, March 1975
pp. 73–80

CIRCULANT MATRICES AND ALGEBRAIC EQUATIONS

ROGER CHALKLEY, University of Cincinnati

1. Introduction. For each monic polynomial

$$f(X) = X^n + c_1X^{n-1} + \cdots + c_n \tag{1}$$

of degree $n \geqq 1$ over the field C of complex numbers, there exist elements $a_1, \cdots, a_n$ in C such that the $n \times n$ circulant matrix

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ . & . & . & \cdots & . \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{bmatrix} \tag{2}$$

has $f(X)$ as its characteristic polynomial in the sense

$$f(X) = \det(XI_n - A). \tag{3}$$

In Section 2, we prove the preceding statement and the identity

$$\det(XI_n - A) = \prod_{s=1}^{n} \left(X - \sum_{k=1}^{n} a_k\zeta_n^{(k-1)(s-1)}\right), \tag{4}$$

where ζ_n denotes a primitive nth root of unity. Thus, the eigenvalues

$$\xi_s = \sum_{k=1}^{n} a_k\zeta_n^{(k-1)(s-1)}, \text{ for } s = 1, 2, \cdots, n, \tag{5}$$

of A are the roots of $f(X)$ in C. *The problem to solve $f(X) = 0$ in C can therefore be replaced by the problem to find an $n \times n$ circulant matrix over C which has $f(X)$ as its characteristic polynomial.* For $n = 1, 2, 3, 4$, all such $n \times n$ circulant matrices are presented in Section 4. There are $n!$ corresponding sets of formulas for the roots of $f(X)$.

2. Properties of circulant matrices. Let n be a positive integer and let F be a field which contains a primitive nth root ζ_n of unity. In particular, the characteristic of F cannot divide n. For $r = 1, \cdots, n$ and $s = 1, \cdots, n$, set $\delta_{rs} = 1$ when $r = s$; and, $\delta_{rs} = 0$ when $r \neq s$.

LEMMA. *We have*

$$\frac{1}{n} \sum_{j=1}^{n} \zeta_n^{(j-1)(s-r)} = \delta_{rs}, \text{ for } r = 1, \cdots, n \text{ and } s = 1, \cdots, n. \tag{6}$$

Proof. Set $\rho = \zeta_n^{(s-r)}$. For $r \neq s$, we obtain $\rho \neq 1$, $\rho^n = 1$, and

$$\frac{1}{n}(1 + \rho + \rho^2 + \cdots + \rho^{n-1}) = \frac{1 - \rho^n}{n(1 - \rho)} = 0.$$

If $r = s$, then $\rho = 1$ and $(1/n)(n) = 1$. This completes the proof.

Let L_n and M_n be the $n \times n$ matrices whose components of row index r and column index s are

$$\lambda_{rs} = \frac{1}{n}\zeta_n^{-(r-1)(s-1)} \text{ and } \mu_{rs} = \zeta_n^{(r-1)(s-1)}.$$

The element in the (r, s) position of $L_n M_n$ is

$$\sum_{j=1}^{n} \lambda_{rj}\mu_{js} = \frac{1}{n} \sum_{j=1}^{n} \zeta_n^{(j-1)(s-r)} = \delta_{rs}.$$

Thus, we have $L_n M_n = I_n$ and $M_n^{-1} = L_n$.

THEOREM. *Suppose $n \times n$ matrices A and D over F satisfy $AM_n = M_n D$. Then, A is a circulant matrix if and only if D is a diagonal matrix.*

Proof. Let a_{rs} and d_{rs} be the components of row index r and column index s for A and D.

(i) Suppose D is a diagonal matrix. With $d_{rs} = d_{rr}\delta_{rs}$ and $A = M_n D M_n^{-1}$, we obtain

$$a_{rs} = \sum_{j=1}^{n} \sum_{k=1}^{n} \mu_{rj} d_{jj} \delta_{jk} \lambda_{ks} = \sum_{j=1}^{n} \mu_{rj} d_{jj} \lambda_{js} = \frac{1}{n} \sum_{j=1}^{n} d_{jj} \zeta_n^{(j-1)(r-s)}.$$

Since the elements in the (r, s) and (r', s') positions of A are equal when $r - s \equiv r' - s'$ $(\text{mod}\, n)$, A is a circulant matrix.

(ii) Suppose A is a circulant matrix. For $r = 1, \cdots, n$ and $s = 1, \cdots, n$, set $a_{r,s+n} = a_{rs}$. With $D = M_n^{-1} A M_n$, we find

$$\begin{aligned} d_{rs} &= \sum_{j=1}^{n} \lambda_{rj} \sum_{k=1}^{n} a_{jk} \zeta_n^{(k-1)(s-1)} = \sum_{j=1}^{n} \lambda_{rj} \sum_{k=j}^{n+j-1} a_{jk} \zeta_n^{(k-1)(s-1)} \\ &= \frac{1}{n} \sum_{j=1}^{n} \zeta_n^{-(r-1)(j-1)} \sum_{k=1}^{n} a_{j,j-1+k} \zeta_n^{(j-1+k-1)(s-1)} \\ &= \frac{1}{n} \sum_{j=1}^{n} \zeta_n^{(j-1)(s-r)} \sum_{k=1}^{n} a_{1k} \zeta_n^{(k-1)(s-1)}. \end{aligned} \tag{7}$$

By (6) and (7), we have $d_{rs} = 0$ when $r \neq s$. This completes the proof.

Corollary 1. *The $n \times n$ circulant matrices over F form a commutative ring under matrix addition and multiplication which is isomorphic to the ring of $n \times n$ diagonal matrices over F.*

Proof. By the theorem, we obtain a one-to-one correspondence between the two sets of matrices which preserves the operations of matrix addition and multiplication. Since the $n \times n$ diagonal matrices over F form a commutative ring, the same is true for the $n \times n$ circulant matrices over F.

Corollary 2. *An $n \times n$ matrix A over F is circulant if and only if all the column vectors of M_n are eigenvectors of A.*

Proof. Set $D = M_n^{-1}AM_n$. From $AM_n = M_nD$, we note that D is a diagonal matrix if and only if each column vector of M_n is an eigenvector of A. To complete the proof, we use the theorem.

Corollary 3. *Let an $n \times n$ circulant matrix A be defined over F by* (2), *and let $\xi_1, \xi_2, \cdots, \xi_n$ be defined by* (5). *Then* $\det A = \xi_1\xi_2\cdots\xi_n$.

Proof. For the diagonal matrix $D = M_n^{-1}AM_n$, we use (7) and (6) with $r = s$ and $a_{1k} = a_k$ to obtain $d_{ss} = \xi_s$. This yields

$$\det A = \det(M_n^{-1}AM_n) = \det D = \prod_{s=1}^{n} \xi_s.$$

For other proofs of this well-known result, see [2] or [1].

Corollary 4. *If A is defined over F by* (2), *then* (4) *is valid.*

Proof. With the notation used for Corollary 3, we obtain

$$\det(XI_n - A) = \det(M_n^{-1}(XI_n - A)M_n) = \det(XI_n - D) = \prod_{s=1}^{n} (X - \xi_s).$$

This completes the proof.

By definition, an $n \times n$ permutation matrix is a matrix obtained from I_n by a permutation of rows (or columns). There are $n!$ such matrices.

Corollary 5. *Suppose $[a_1, \cdots, a_n]$ and $[b_1, \cdots, b_n]$ are the first rows of $n \times n$ circulant matrices A and B over F. Then, A and B have the same characteristic polynomial if and only if there exists an $n \times n$ permutation matrix P such that*

$$[b_1, \cdots, b_n] = [a_1, \cdots, a_n](M_nPM_n^{-1}). \tag{8}$$

Proof. Define $\xi_1, \cdots, \xi_n$ and $\eta_1, \cdots, \eta_n$ by

$$[\xi_1, \cdots, \xi_n] = [a_1, \cdots, a_n]M_n \text{ and } [\eta_1, \cdots, \eta_n] = [b_1, \cdots, b_n]M_n.$$

We use Corollary 4 to obtain

$$\det(XI_n - A) = \prod_{s=1}^{n} (X - \xi_s) \text{ and } \det(XI_n - B) = \prod_{s=1}^{n} (X - \eta_s).$$

Thus, the characteristic polynomials of A and B are equal if and only if there exists an $n \times n$ permutation matrix P such that

$$[\eta_1, \cdots, \eta_n] = [\xi_1, \cdots, \xi_n]P. \tag{9}$$

To complete the proof, we rewrite (9) in the form (8).

COROLLARY 6. *Suppose A is an $n \times n$ circulant matrix over F. Then, an $n \times n$ matrix B over F is circulant with*

$$\det(XI_n - B) = \det(XI_n - A) \tag{10}$$

if and only if there exists an $n \times n$ permutation matrix P such that

$$B = (M_n P^T M_n^{-1}) A (M_n P M_n^{-1}). \tag{11}$$

Proof. (i) Suppose B is circulant and related to A by (10). Then, the characteristic polynomials of the diagonal matrices $D = M_n^{-1} A M_n$ and $E = M_n^{-1} B M_n$ are equal. Hence, there exists an $n \times n$ permutation matrix P such that $E = P^T D P$. This yields (11).

(ii) Suppose (11) is satisfied. Then, we easily obtain (10). Moreover, $M_n^{-1} A M_n$ is diagonal; $P^T(M_n^{-1} A M_n)P$ is diagonal; and, with (11), B is circulant. This completes the proof.

Set $R_n = (1/n)M_n^2$. The element in the (r, s) position of R_n is

$$\frac{1}{n} \sum_{j=1}^{n} \mu_{rj}\mu_{js} = \frac{1}{n} \sum_{j=1}^{n} \zeta_n^{(j-1)(r+s-2)};$$

it equals 1 when n divides $r + s - 2$ and it equals 0 otherwise. Thus, R_n is a symmetric permutation matrix. To obtain R_n from I_n, we can interchange the jth and $(n + 2 - j)$th rows of I_n for $j = 2, 3, \cdots, n$. Also, R_n results when the jth and $(n + 2 - j)$th columns of I_n are interchanged for $j = 2, 3, \cdots, n$. With $R_n^2 = I_n$, we have $M_n^4 = n^2 I_n$ and $M_n^{-1} = (1/n^2)M_n^3$.

Suppose A is given by (2). We can obtain the transpose A^T of A as follows: for $j = 2, 3, \cdots, n$, interchange the jth and $(n + 2 - j)$th rows of A; then, for $j = 2, 3, \cdots, n$, interchange the jth and $(n + 2 - j)$th columns of the resulting matrix. Thus, we have $R_n A R_n = A^T$.

Suppose $n \times n$ matrices P and Q over F satisfy $Q = R_n P R_n$. We note that Q is circulant if and only if P is circulant; Q is diagonal if and only if P is diagonal; and, Q is a permutation matrix if and only if P is a permutation matrix. The relations

$$M_n P M_n^{-1} = M_n^{-1} Q M_n \text{ and } M_n P^T M_n^{-1} = M_n^{-1} Q^T M_n$$

can be used to reformulate (8) and (11).

COROLLARY 7. *Let A and D be $n \times n$ matrices over F such that $AM_n = M_nD$. Then, one of A and D is circulant if and only if the other is diagonal.*

Proof. We have $M_nAM_n^{-1} = M_n^2DM_n^{-2} = R_nDR_n$. By the theorem, A is diagonal if and only if $M_nAM_n^{-1}$ is circulant; but, R_nDR_n is circulant if and only if D is circulant. Directly from the theorem, A is circulant if and only if D is diagonal. This completes the proof.

Henceforth, we specialize F to be the field C of complex numbers.

COROLLARY 8. *If $f(X)$ is a monic polynomial of degree $n \geqq 1$ over C, then $f(X)$ is the characteristic polynomial of some $n \times n$ circulant matrix over C.*

Proof. There exist elements $\eta_1, \eta_2, \cdots, \eta_n$ in C such that

$$f(X) = (X - \eta_1)(X - \eta_2)\cdots(X - \eta_n).$$

Let D be the $n \times n$ diagonal matrix with $d_{ss} = \eta_s$. Set $A = M_nDM_n^{-1}$. Then, A is an $n \times n$ circulant matrix over C and

$$\det(XI_n - A) = \det(M_n^{-1}(XI_n - A)M_n) = \det(XI_n - D) = f(X).$$

This completes the proof.

Set $N_n = (1/\sqrt{n})M_n$. Since the complex conjugate of ζ_n^k is ζ_n^{-k}, the conjugate $\bar{N}_n$ of N_n satisfies $\bar{N}_n = N_n^{-1}$. With $N_n^T = N_n$ and $\bar{N}_n^T = N_n^{-1}$, the matrix N_n is both symmetric and unitary. We note that $N_n^2 = R_n$ and $N_n^{-1} = N_n^3$. When $n \geqq 3$, N_n generates a cyclic group of order 4. For each $n \times n$ matrix P, we have $M_nPM_n^{-1} = N_nPN_n^{-1}$.

3. Several group representations. The elements of the symmetric group S_n are the permutations of n objects. When the objects are identified with the rows of I_n, we obtain an isomorphism of S_n onto the multiplicative group of $n \times n$ permutation matrices. As P ranges over the $n \times n$ permutation matrices, the corresponding matrices $M_nPM_n^{-1}$ also form a group under matrix multiplication which is isomorphic to S_n.

LEMMA. *The element in the $(1,1)$ position of $M_nPM_n^{-1}$ equals 1 and the elements in the other positions of the first row or first column equal 0.*

Proof. For $P = [\pi_{rs}]$ and $M_nPM_n^{-1} = [\alpha_{rs}]$, we find

$$\alpha_{rs} = \frac{1}{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \zeta_n^{(r-1)(j-1)} \pi_{jk} \zeta_n^{-(k-1)(s-1)}$$

and

$$\alpha_{1s} = \frac{1}{n} \sum_{k=1}^{n} \zeta_n^{(k-1)(1-s)} \sum_{j=1}^{n} \pi_{jk} = \delta_{1s}.$$

Similarly, we obtain $\alpha_{r1} = \delta_{r1}$. This completes the proof.

PROPOSITION. *For $n \geqq 2$, let P range over all $n!$ of the permutation matrices of size $n \times n$. Then, the corresponding matrices of size $(n-1)\times(n-1)$ obtained by deletion of the first row and first column from each matrix $M_n P M_n^{-1}$ form a group under matrix multiplication which is isomorphic to S_n.*

Proof. This follows directly from the lemma.

Example 1. For $n = 3$, $\omega^2+\omega+1=0$, $\zeta_3=\omega$,

$$M_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix}, \text{ and } 3M_3^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega \\ 1 & \omega & \omega^2 \end{bmatrix},$$

we follow the details of the proposition to obtain

$$(12)\quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \omega & 0 \\ 0 & \omega^2 \end{bmatrix}, \begin{bmatrix} \omega^2 & 0 \\ 0 & \omega \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & \omega^2 \\ \omega & 0 \end{bmatrix}, \begin{bmatrix} 0 & \omega \\ \omega^2 & 0 \end{bmatrix}.$$

The six matrices of (12) form a group isomorphic to S_3.

The lemma and the proposition remain valid when M_n is replaced throughout by $G_n M_n$, where $G_n = [g_{rs}]$ is a nonsingular $n \times n$ matrix with

$$g_{r1} = g_{1r} = \delta_{r1}, \text{ for } r = 1, 2, \cdots, n.$$

For instance, set

$$g_{rs} = -\zeta_n^{-(r-1)(s-1)}, \text{ for } r = 2, 3, \cdots, n \text{ and } s = 2, 3, \cdots, n,$$

$H_n = G_n M_n$, and $H_n = [h_{rs}]$. We find $h_{rr} = 1-n$ for $r = 2, 3, \cdots, n$ and $h_{rs} = 1$ otherwise; moreover,

$$nH_n^{-1} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix}.$$

Example 2. For $n = 3$, the matrices $H_3 P H_3^{-1}$ yield

$$(13)\quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}.$$

The six matrices of (13) form a group which is isomorphic to S_3. For related details, see [3].

The even permutations in S_n correspond to matrices P, $M_n P M_n^{-1}$, $H_n P H_n^{-1}$, etc., with determinant equal to 1; and, the odd permutations in S_n correspond to matrices with determinant equal to -1. For example, the alternating subgroup of S_3 is represented by the first three matrices in (13).

4. The roots of $f(X)$, for $n = 1,2,3,4$. First, we relate $c_1, \cdots, c_n$ to $a_1, \cdots, a_n$ in (1), (2), and (3). From (1), the $(n-k)$th derivative $f^{(n-k)}(X)$ of $f(X)$ yields

$$c_k = \frac{f^{(n-k)}(0)}{(n-k)!}, \quad \text{for } k = 1, 2, \cdots, n.$$

By application to (3) of rules for the differentiation of a determinant, we conclude: *c_k equals the sum of the determinants of the principal $k \times k$ submatrices of $-A$.*

When $n = 1$, we find $\zeta_1 = 1$, $c_1 = -a_1$, and $\xi_1 = a_1 = -c_1$.

For $n = 2$, we need $\zeta_2 = -1$, $c_1 = -2a_1$, $c_2 = a_1^2 - a_2^2$, $a_1 = -c_1/2$, and a_2 such that $a_2^2 = (c_1^2/4) - c_2$; by (5), the roots of $f(X)$ in C are $a_1 + a_2$ and $a_1 - a_2$.

In each case, we have $c_1 = -na_1$. For $n = 3$ and $n = 4$, it is convenient to make a preliminary transformation so that $c_1 = 0$.

A 3×3 circulant matrix A given by (2) with $n = 3$ has

$$\det(XI_3 - A) = X^3 + c_2X + c_3 \tag{14}$$

as its characteristic polynomial if and only if its components satisfy

$$-3a_1 = 0, \quad -3a_2a_3 = c_2, \text{ and } -a_2^3 - a_3^3 = c_3. \tag{15}$$

When $c_2 = c_3 = 0$, set $y_0 = z_0 = 0$; otherwise, $T^2 + c_3T - (c_2/3)^3$ has a nonzero root t_0, let y_0 satisfy $Y^3 = t_0$, and set $z_0 = (-c_2)/(3y_0)$. In this way, (15) is satisfied with $a_1 = 0$, $a_2 = y_0$, $a_3 = z_0$. We use (8) and the six matrices $M_3PM_3^{-1}$ upon which (12) was based to obtain the first rows

$$\begin{gathered}[0, y_0, z_0],\ [0, y_0\omega, z_0\omega^2],\ [0, y_0\omega^2, z_0\omega],\\ [0, z_0, y_0],\ [0, z_0\omega, y_0\omega^2],\ [0, z_0\omega^2, y_0\omega]\end{gathered} \tag{16}$$

of all 3×3 circulant matrices (2) for (14). Now, we can use (5) to write all six forms of Cardan's formulas for the roots of a cubic. Namely, for each selection of $[a_1, a_2, a_3]$ from (16), the corresponding elements

$$a_2\omega^{(s-1)} + a_3\omega^{2(s-1)}, \text{ for } s = 1, 2, 3,$$

are the roots in C of $X^3 + c_2X + c_3$.

A 4×4 circulant matrix A given by (2) with $n = 4$ has

$$\det(XI_4 - A) = X^4 + c_2X^2 + c_3X + c_4 \tag{17}$$

as its characteristic polynomial if and only if its components satisfy

$$\begin{gathered}-4a_1 = 0, \quad -4a_2a_4 - 2a_3^2 = c_2, \quad -4a_2^2a_3 - 4a_3a_4^2 = c_3, \text{ and}\\ -a_2^4 + a_3^4 - a_4^4 + 2a_2^2a_4^2 - 4a_2a_3a_4^2 = c_4.\end{gathered} \tag{18}$$

When $c_2 = c_3 = c_4 = 0$, set $u_0 = v_0 = w_0 = 0$; otherwise, the equation

$$(4V^2)^3 + 2c_2(4V^2)^2 + (c_2^2 - 4c_4)(4V^2) - c_3^2 = 0$$

has a nonzero solution v_0 and we select u_0, w_0 to satisfy

$$UW = -\frac{v_0^2}{2} - \frac{c_2}{4} \text{ and } U^2 + W^2 = -\frac{c_3}{4v_0}.$$

We can verify that $a_1 = 0$, $a_2 = u_0$, $a_3 = v_0$, $a_4 = w_0$ is a solution of (18); similar details were given in [1]. Next, we apply (8). As P ranges over the permutation matrices of size 4×4, the formula

$$[a_1, a_2, a_3, a_4] = [0, u_0, v_0, w_0](M_4 P M_4^{-1})$$

specifies the first row $[a_1, a_2, a_3, a_4]$ of each 4×4 circulant matrix A for (17). With $i^2 = -1$ and $\zeta_4 = i$, we use (5) to conclude that the corresponding elements

$$a_2 i^{(s-1)} + a_3 i^{2(s-1)} + a_4 i^{3(s-1)}, \quad \text{for } s = 1, 2, 3, 4,$$

are the roots in C of $X^4 + c_2 X^2 + c_3 X + c_4$. Thus, there are $4! = 24$ sets of solution formulas for a biquadratic analogous to the $3! = 6$ forms of Cardan's formulas for a cubic.

References

1. R. Chalkley, Cardan's formulas and biquadratic equations, this MAGAZINE, 47 (1974) 8–14.
2. T. Muir, A Treatise on the Theory of Determinants, Dover, New York, 1960, pp. 442–445.
3. E. D. Nering, Linear Algebra and Matrix Theory, Wiley, New York, 1970, pp. 306–308.