# ROOTS AND CANONICAL FORMS FOR CIRCULANT MATRICES

BY

C. M. ABLOW AND J. L. BRENNER

1. **Introduction.** A square matrix is called circulant([1]) if each row after the first is obtained from its predecessor by a cyclic shift. Circulant matrices arise in the study of periodic or multiply symmetric dynamical systems. In particular they have application in the theory of crystal structure [1].

The history of circulant matrices is a long one. In this paper a (block-diagonal) canonical form for circulant matrices is derived. The matrix which transforms a circulant matrix to canonical form is given explicitly. Thus the characteristic roots and vectors of the original circulant can be found by solving matrices of lower order.

If the cyclic shift defining the circulant is a shift by one column([2]) to the right, the circulant is called simple. Many of the theorems demonstrated here are well known for simple circulants. The theory has been extended to general circulant and composite circulant matrices by B. Friedman [3]. The present proofs are different from his; some of the results obtained go beyond his work.

2. **Notations.**

DEFINITION 2.1. *A g-circulant matrix is an $n \times n$ square matrix of complex numbers, in which each row (except the first) is obtained from the preceding row by shifting the elements cyclically g columns to the right.*

This connection between the elements $a_{ij}$ of the $i$th row and the elements of the preceding row is repeated in the formula

$$a_{ij} = a_{i-1,\, j-g}, \tag{2.1}$$

where indices are reduced to their least positive remainders modulo $n$. If equation (2.1) holds for all values of $i$ greater than 1, it will hold automatically for $i = 1$.

It is possible to generalize the methods and results of this paper by allowing the elements $a_{ij}$ of the circulant matrix to be square matrices themselves, all of fixed dimension. This extension is outlined in §6 below.

Let $A$ be an arbitrary matrix. If there is a nonzero vector $x$ and a scalar $\lambda$ such that the relation

Presented to the Society April 23, 1960 under the title *Circulant and composite circulant matrices*; received by the editors November 13, 1961.

(1) Rutherford [5] uses the term continuant for circulant.

(2) See the example of a 5-circulant on p. 31.

$$Ax = x\lambda \tag{2.2}$$

holds, then $\lambda$ is called a characteristic root (proper value, eigenvalue) of $A$, and $x$ a corresponding vector. There may be several vectors corresponding to the same root, but no more than one root corresponding to the same vector, for a fixed matrix $A$. The significant properties of a matrix are all known when its vectors, roots, and invariant spaces are found(3). The process of finding these is called "solving the matrix." The general circulant matrix is solved in this article.

The chief tool used in solving the matrix $A$ is the relation $PA = AP^g$ which is established in Theorem 2.1. In this relation $P$ is a certain permutation matrix. This relation is effective because all the roots and vectors of $P$ can be given.

DEFINITION 2.2. *$P_n$ is the $n \times n$ 1-circulant*

$$\begin{bmatrix} 0, & 1, & 0, & \cdots, & 0 \\ 0, & 0, & 1, & \cdots, & 0 \\ & & \vdots & & \\ 1, & 0, & 0, & \cdots, & 0 \end{bmatrix}. \tag{2.3}$$

LEMMA 2.1. *Let $\omega = \exp\{2\pi i/n\}$, a primitive nth root of unity, and let $x(h)$ be the n-vector $[1, \omega^h, \omega^{2h}, \cdots, \omega^{(n-1)h}]'$, a column of n numbers(4). The various powers of $\omega$, $\omega^h$, are proper values of $P_n$ and the $x(h)$ corresponding vectors:*

$$P_n x(h) = x(h)\omega^h \quad (h = 1, 2, \cdots, n). \tag{2.4}$$

Equations (2.4) may be verified directly. Since the proper values of $P_n$ are distinct, the corresponding vectors are linearly independent. Thus the matrix $M$, whose $h$th column is $x(h)$, is nonsingular. Combining (2.4) into a single matrix equation gives

$$P_n M = M \operatorname{diag}[\omega, \omega^2, \cdots, \omega^{n-1}, 1]. \tag{2.5}$$

From this $M^{-1}P_n M = \operatorname{diag}[\cdots]$ which solves $P_n$.

THEOREM 2.1. *The equation*

$$P_n A = AP_n^g \tag{2.6}$$

*characterizes the g-circulant property of A. That is, the matrix A is a g-circulant matrix if and only if relation (2.6) is valid.*

(3) An invariant space belonging to $A$ is a set of vectors $M$, closed under addition and multiplication by scalars, such that $Ax$ is a member of $M$ whenever $x$ is in $M$. In modern words, $M$ is a linear manifold which admits $A$. In older terminology, solving a matrix $A$ means finding its Jordan canonical form, J, and a matrix $N$ which transforms $A$ into $J : N^{-1}AN = J$. The diagonal blocks of $J$ together with corresponding columns of $N$ exhibit the vectors, roots, and the invariant spaces of $A$.

(4) The prime denotes the transpose operation.

**Proof.** The matrix $P_nA$ is obtained from the matrix $A$ by raising each row of $A$ and placing the first row of $A$ at the bottom. On the other hand, the matrix $AP_n$ is obtained from the matrix $A$ by permuting each row cyclically, so that $AP_n^g$ is obtained from $A$ by $g$ such cyclic permutations. The theorem follows.

3. **General theorems.** The general theorems of this section seem to be new. They are easily established from Theorem 2.1, and are used in turn to decompose a circulant matrix into block-diagonal form. At the end of the section, a recent theorem of Lewis [4] is rederived.

THEOREM 3.1. *If $A$ is a $g$-circulant and $B$ is an $h$-circulant, then $AB$ is a $gh$-circulant.*

The first step in the proof is to establish the formula

$$P_n^gB = BP_n^{gh}$$

by induction from the formula $P_nB = BP_n^h$ which is implied by the hypothesis. The proof is completed by using the other part of the hypothesis, $P_nA = AP_n^g$, to derive the equalities

$$P_nAB = AP_n^gB = ABP_n^{gh}.$$

THEOREM 3.2. *Let $h$ be an integer, $1 \leqq h \leqq n$. If $A$ is a $g$-circulant, there is a scalar $w(h, A)$ such that*

$$Ax(h) = x(hg)w(h, A). \tag{3.1}$$

This theorem states that $A$ carries one vector of $P_n$ into some fixed multiple of another vector of $P_n$ (possibly the same one). The proof uses Theorem 2.1. First one establishes the formula

$$P_n^gx(h) = x(h)\omega^{gh} \tag{3.2}$$

by induction from (2.4). From (2.6) and (3.2) one concludes that

$$P_n\{Ax(h)\} = \{Ax(h)\}\omega^{gh}, \tag{3.3}$$

and (3.1) follows from this and from the additional remark that every vector $y$ satisfying $P_ny = y\omega^{gh}$ must be a multiple of $x(gh)$.

When $A$ is a 1-circulant (classical circulant), $g = 1$ and (3.1) exhibits proper values and corresponding vectors of $A$. The solution of $A$ is obtained at once.

THEOREM 3.3. *Let $A$ be a 1-circulant, and let $M$ be the $n \times n$ matrix with $h$th column $x(h)$: $M = [x(1), x(2), \cdots, x(n)]$. Then*

$$M^{-1}AM = \operatorname{diag}[w(1, A), w(2, A), \cdots, w(n, A)] = D.$$

The reader should note that this decomposition (solution) of the matrix $A$ is effective, since equation (3.1) actually provides a formula for the quantity $w(h, A)$. This is so because the first element of $x(gh)$ is 1, so that $w(h, A)$ is the (well-defined) first element of $Ax(h)$:

$$w(h, A) = a_{11} + a_{12}\omega^h + \cdots + a_{1n}\omega^{(n-1)h}.$$

A theorem of Lewis [4] is corollary to the above results. The theorem asserts that, if $A$ is 1-circulant, det $A$ is a symmetric function of the elements of $A$ if and only if $A$ is of order 1 or 3. A short proof is the following.

If $A$ is of order $n \geqq 3$, $\det A = w(1, A)w(2, A) \cdots w(n, A)$, as is evident from Theorem 3.3. This factorization of $\det A$ is unique for polynomials in the indeterminates $a_i$, the elements of the first row of $A$. Thus if $\det A$ is to be unaffected by the interchange of $a_2$ and $a_3$ say, then each $w(h, A)$ must be mapped by the interchange into $w(h', A)$ for some $h'$. In symbols,

$$a_1 + a_3\omega^h + a_2\omega^{2h} + \cdots = a_1 + a_2\omega^{h'} + a_3\omega^{2h'} + \cdots,$$

whence

$$h \equiv 2h' \equiv 2(2h) \bmod n$$

for every $h$. If $n \geqq 3$, this implies $n = 3$. The assertions for $n < 3$ are subject to simple verification.

4. **Prime circulants.** The solution of a $g$-circulant matrix offers special difficulties if $g$ and $n$ have a common factor greater than unity. In this section, we show how to handle the case where $g$ and $n$ are relatively prime; in the next section, we take up the case $g = 0$, and finally in §8, a method is developed for the general case where $g$ and $n$ have a common factor between 1 and $n$. The method of §8 requires results on circulants, the elements of which are themselves matrices. These results are natural generalizations of the results of §§2–5; the proofs of the general results are obtained by a natural extension principle, as will be indicated in §6.

LEMMA 4.1. *If $A$ is a $g$-circulant, the relation*

$$w(h, A^k) = w(g^{k-1}h, A)w(g^{k-2}h, A) \cdots w(gh, A)w(h, A) \tag{4.1}$$

*holds.*

**Proof.** From (3.1) the relation $Ax(g^t h) = x(g^{t+1}h)w(g^t h, A)$ follows. From this, one obtains by induction the relation

$$A^k x(h) = x(g^k h)w(g^{k-1}h, A)w(g^{k-2}h, A) \cdots w(gh, A)w(h, A). \tag{4.2}$$

On the other hand, Theorem 3.1 shows that $A^k$ is a $g^k$-circulant, so that from (3.1) one also obtains the relation $A^k x(h) = x(g^k h)w(h, A^k)$. Combining this result with (4.2), the assertion of the lemma is obtained.

The following definition gives an equivalence relation (introduced by Friedman [3]) on which the solution of a $g$-circulant matrix depends.

DEFINITION 4.1. *Let $(g,n) = 1$. The equivalence relation "$\sim$" on the residue classes* $1, 2, \cdots, n$ mod $n$ *is defined as follows*:

$$h_1 \sim h_2 \leftrightarrow \exists q, \quad h_1 \equiv h_2 g^q (\mathrm{mod}\, n).$$

Thus $h_1$, $h_2$ are equivalent if one arises from the other on multiplication by a positive power of $g$. This definition is obviously reflexive and transitive; it is symmetric because of Euler's generalization of Fermat's little theorem: $g^{\phi(n)} \equiv 1 \,(\mathrm{mod}\, n)$. Thus $h_2 \equiv h_1 g^{q\phi(n)-q} \;(\mathrm{mod}\, n)$.

Since "$\sim$" is an equivalence relation, it separates the residue classes $1, 2, \cdots, n$ into equivalence classes (mutually exclusive and exhaustive). The class to which $h$ belongs is denoted by $C(h, g, n)$; it consists of the numbers $h, hg, hg^2, \cdots, hg^{f-1}$ $(\mathrm{mod}\, n)$, where $f$ is the smallest exponent for which the relation

$$hg^f \equiv h \,(\mathrm{mod}\, n) \tag{4.3}$$

holds. One sees that $f$ is the index to which $g$ belongs mod $\{n/(h,n)\}$.

The next theorem gives a block diagonal form of a $g$-circulant matrix. It is known that the roots and vectors of a block diagonal matrix can be found by solving the blocks individually (as lower order matrices). Thus, Theorem 4.1 reduces the problem of solving $A$ to the problem of solving matrices of lower order. The matrices of lower order are then solved explicitly.

THEOREM 4.1. *Let $A$ be an $n \times n$ $g$-circulant matrix, $(g,n) = 1$. Let $\{C(h_i, g, n)\}$ $(i = 1, 2, \cdots, t)$ be a complete set of equivalence classes under the equivalence "$\sim$", where the ith class has $f_i$ elements. Thus $h_1, h_2, \cdots, h_t$ forms a complete set of representatives of these equivalence classes. The block-diagonal form of $A$ is given by*

$$N^{-1}AN = \mathrm{diag}\,[W(h_1, A), W(h_2, A), \cdots, W(h_t, A)] = D_1, \tag{4.4}$$

*say, where*

$$\begin{aligned} N &= [X(h_1), X(h_2), \cdots, X(h_t)], \\ X(h_i) &= [x(h_i), x(gh_i), \cdots, x(g^{f_i-1}h_i)], \\ W(h_i, A) &= P_{f_i}^{-1}\,\mathrm{diag}\,[w(h_i, A), w(gh_i, A), \cdots, w(g^{f_i-1}h_i, A)]. \end{aligned} \tag{4.5}$$

In the statements of this theorem, $W(h_i, A)$ is an $f_i \times f_i$ matrix (called a broken diagonal matrix by Friedman [2; 3]); $X(h_i)$ is a matrix with $f_i$ columns and $n$ rows; $N$ is the $n \times n$ matrix obtained by writing the matrices $X(h_i)$ one beside the other. Since the columns of $N$ are the vectors of $P_n$ in a particular order, it is clear that $N$ is invertible. The matrix $W(h, A)$ has in fact the form

$$\text{(4.6)} \qquad \begin{bmatrix} 0, & 0, & \cdots, & 0, & w(g^{f-1}h, A) \\ w(h, A), & 0, & \cdots, & 0, & 0 \\ 0, & w(gh, A), & \cdots, & 0, & 0 \\ & & \cdot \ \cdot \ \cdot & & \end{bmatrix}.$$

Theorem 4.1 is essentially a restatement of Theorem 3.2, using the notation of Definition 4.1 and the remark embodied in congruence (4.3). Hence Theorem 4.1 requires no proof, but only verification of the relation $AN = ND_1$. This amounts to a series of equations, of which a typical set is (see Theorem 3.2)

$$\text{(4.7)} \qquad \begin{aligned} Ax(h_i) &= x(gh_i)w(h_i, A), \\ Ax(gh_i) &= x(g^2h_i)w(gh_i, A), \\ &\cdots \\ Ax(g^{f_i-1}h_i) &= x(h_i)w(g^{f_i-1}h_i, A). \end{aligned}$$

LEMMA 4.2. *Let $\mu$ be a root of $W(h_i, A)$, and $v$ a corresponding vector. Then $\mu$ is a root of $A$, and $X(h_i)v$ is a corresponding vector. Moreover, all roots and vectors of $A$ arise in this way.*

This lemma is also subject to direct verification. Thus a complete solution of $A$ is obtained from the following sequence of lemmas, which show how to solve a typical matrix $W(h, A)$.

LEMMA 4.3. *Let $a_f a_{f-1} \cdots a_1 \neq 0$. Then the roots of the $f \times f$ matrix*

$$\text{(4.8)} \qquad W = \begin{bmatrix} 0, & 0, & \cdots, & a_f \\ a_1, & 0, & \cdots, & 0 \\ 0, & a_2, & \cdots, & 0 \\ & & \cdot \ \cdot \ \cdot & \end{bmatrix}$$

*are the $f$th roots of $a_f a_{f-1} \cdots a_1$, and a vector corresponding to the root $\lambda$ is*

$$[\lambda^{f-1}, a_1\lambda^{f-2}, a_2a_1\lambda^{f-3}, \cdots, a_{f-1}a_{f-2} \cdots a_1]'.$$

This lemma is easily verified directly.

The following discussion is concerned with the case $a_f a_{f-1} \cdots a_1 = 0$.

LEMMA 4.4. *Let $W$ be the matrix (4.8). Let $a_r = 0$ and $a_{r+1}a_{r+2} \cdots a_{r+k} \neq 0$. Let $R_{r,k+1} = (E_{ij})$ be the rectangular matrix of $(k+1)$ columns and $f$ rows with all elements zero except for the following*

$$E_{1r} = 1,\ E_{2,r+1} = a_{r+1}, \cdots, E_{k,r+k-1} = a_{r+k}a_{r+k-1} \cdots a_{r+1}.$$

*Then*

$$WR_{r,k+1} = R_{r,k+1}H_{k+1},$$

*where $H_s$ is the square matrix of order s with all zeros except for* 1*'s in its main subdiagonal:*

$$(4.9) \qquad H_s = \begin{bmatrix} 0, & 0, & \cdots, & 0, & 0 \\ 1, & 0, & \cdots, & 0, & 0 \\ 0, & 1, & \cdots, & 0, & 0 \\ \cdot & \cdot & \cdots, & \cdot & \cdot \\ 0, & 0, & \cdots, & 1, & 0 \end{bmatrix}, \quad H_1 = [0].$$

The result may be verified directly.

LEMMA 4.5. *Let $W$ be the matrix* (4.8) *and let $[R_{r_1,t_1}, \cdots, R_{r_p,t_p}]$ be a complete set for $W$ of the matrices $R_{r,k+1}$ introduced in Lemma* 4.4, *each $R_{r,t}$ being of maximal size. Then*

$$N = [R_{r_1,t_1}, R_{r_2,t_2}, \cdots, R_{r_p,t_p}]$$

*is nonsingular and*

$$WN = N \operatorname{diag}[H_{t_1}, H_{t_2}, \cdots, H_{t_p}].$$

That $R_{r,t}$ is of maximal size means that $a_r = 0, a_{r+1}a_{r+2} \cdots a_{r+t-1} \neq 0$, and $a_{r+t} = 0$. Thus if the $R_{r_i,t_i}$ in $N$ are in proper order, a certain complete subdiagonal of $N$ will have all nonzero elements while all other elements of $N$ are zero. It follows that $N$ is nonsingular as needed. Of course, $\operatorname{diag}[H_{t_1}, \cdots, H_{t_p}]$ is a Jordan form for $W$ so that $W$ has been solved.

5. **Zero circulants.** If $g = 0$, $n$th order matrix $A$ satisfies $P_nA = A$ and all rows of $A$ are the same. If $r$ is the row vector formed from elements of a row of $A$, then

$$(5.1) \qquad A = x(n)r, \; x(n) = [1, 1, \cdots, 1]'.$$

If $x$ is a characteristic (column) vector of $A$ corresponding to characteristic value $\lambda$ then

$$Ax = x(n)rx = x\lambda.$$

If $\lambda \neq 0$, since $rx$ is a scalar, $x$ is proportional to $x(n)$ and $\lambda = rx(n) = w(n, A)$. If $\lambda = 0$, $x$ is a solution of $rx = 0$. Assembling $x(n)$ and any $(n - 1)$ linearly independent solution vectors of $rx = 0$ to form the columns of matrix $N$, one obtains

$$AN = N \operatorname{diag}[w(n, A), 0, 0, \cdots, 0]$$

with nonsingular $N$; this solves $A$.

The same solution of $A$ is valid if $A$ is the zero matrix.

If $A$ is not identically zero but $w(n, A) = 0$, assemble nonsingular matrix $N$ from the column vectors $x(n)$, $r'/rr'$, and $(n-2)$ solution vectors of $rx = 0$ linearly independent of $x(n)$ and each other. Then one may readily verify by (5.1) that

$$AN = NJ$$

where Jordan matrix $J$ is zero except for a unit element in the first row, second column.

6. **Composite circulants.** The solution of $n$th order $g$-circulant matrices in case $g$ and $n$ have common factors between 1 and $n$ can be reduced to the case of zero circulant composite matrices, a composite matrix being a matrix whose elements are themselves matrices. It is therefore expedient to inquire to what extent previous theorems apply to composite matrices.

Unless indicated otherwise, the composite matrices considered are square matrices of order $n$ with square submatrices of order $m$ as elements. Composite matrices are indicated by bold-face type.

Matrix $\mathbf{P}_n$ is the composiite matrix of the form (2.3) with the zero elements of that form replaced by zero matrices of order $m$ and the units replaced by unit matrices of order $m$. The analogues of equations (2.4),

$$\mathbf{P}_n\mathbf{x}(h) = \mathbf{x}(h)\boldsymbol{\omega}^h \qquad (h = 1, 2, \cdots, n),$$

are valid with $\boldsymbol{\omega}^h$ the scalar matrix of order $m$, i.e., $\boldsymbol{\omega}^h$ times the identity matrix, and $\mathbf{x}(h) = [\mathbf{1}, \boldsymbol{\omega}^h, \boldsymbol{\omega}^{2h}, \cdots, \boldsymbol{\omega}^{(h-1)h}]'$. The columns of $\mathbf{x}(h)$ are seen to span the invariant subspace of $\mathbf{P}_n$ corresponding to characteristic value $\omega^h$. Since these columns are linearly independent and independent of the columns of similar composite vectors from other subspaces of $\mathbf{P}_n$, composite matrix $\mathbf{M}$, whose $h$th composite column is $\mathbf{x}(h)$, is nonsingular, and the analogue of (2.5) holds

$$\mathbf{M}^{-1}\mathbf{P}_n\mathbf{M} = \operatorname{diag}[\boldsymbol{\omega}, \boldsymbol{\omega}^2, \cdots, \boldsymbol{\omega}^{n-1}, \mathbf{1}].$$

The composite analogues of Theorems 2.1 to 3.3 are established by a mere reinterpretation of the various steps of the proofs. For example, the critical step in the proof of the analogue of Theorem 3.2 would now be stated as follows. From

$$\mathbf{P}_n\{\mathbf{A}\mathbf{x}(h)\} = \{\mathbf{A}\mathbf{x}(h)\}\boldsymbol{\omega}^{gh}$$

one sees that composite vector $\mathbf{A}\mathbf{x}(h)$ is in the invariant subspace of $\mathbf{P}_n$ corresponding to characteristic value $\omega^{gh}$. There exists therefore an $m \times m$ matrix $\mathbf{w}(gh, \mathbf{A})$ exhibiting the dependence of the columns of $\mathbf{A}\mathbf{x}(h)$ on the basis vectors in that subspace, the columns of $\mathbf{x}(gh)$:

$$\mathbf{A}\mathbf{x}(h) = \mathbf{x}(gh)\mathbf{w}(gh, \mathbf{A}).$$

As before $\mathbf{w}(gh, \mathbf{A})$ is constructively given as the first submatrix element of composite vector $\mathbf{Ax}(h)$.

The conclusion of Theorem 3.3 now shows that the solution of composite 1-circulant $\mathbf{A}$ is obtained by solving the simple matrices $\mathbf{w}(h, \mathbf{A})$. For if $\mathbf{N}_h$ transforms $\mathbf{w}(h, \mathbf{A})$ into Jordan form then $\mathbf{M}\,\mathrm{diag}\,[\mathbf{N}_1, \mathbf{N}_2, \cdots, \mathbf{N}_n]$ will so transform $\mathbf{A}$.

No difficulty arises in the extension to composite matrices of the definitions, proofs, and results of §4 leading to Theorem 4.1. That theorem shows that in the composite case the solution of prime circulant $\mathbf{A}$ has been reduced to the solution of composite matrices $\mathbf{W}(h_i, \mathbf{A})$ of lower order. The result of Lemma 4.2 that roots and vectors of $\mathbf{A}$ are obtained from those of the various $\mathbf{W}(h_i, \mathbf{A})$ is also valid. But the detailed decomposition of $W(h_i, A)$ given in the lemmas following cannot be carried over directly to the composite case.

In the condensed notation of Lemma 4.3, let $\mathbf{a}_i$ stand for $\mathbf{w}(g^{i-1}h, \mathbf{A})$ and $\mathbf{a}$ for $\mathbf{w}(h, \mathbf{A}^f) = \mathbf{a}_f \mathbf{a}_{f-1} \cdots \mathbf{a}_1$. Then $\mathbf{W}(h, \mathbf{A}) = \mathbf{P}_f^{-1}\ \mathrm{diag}\,(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_f)$ and $[\mathbf{W}(h, \mathbf{A})]^f = \mathrm{diag}\,(\mathbf{a}, \mathbf{a}, \cdots, \mathbf{a})$. Let $\boldsymbol{\rho}$ be a matrix of $m$th order transforming $\mathbf{a}$ into Jordan form:

$$\boldsymbol{\rho}^{-1}\mathbf{a}\boldsymbol{\rho} = \mathrm{diag}\,(\mathbf{j}_1, \mathbf{j}_0)$$

where $\mathbf{j}_1$ is nonsingular and $\mathbf{j}_0$ is nilpotent, i.e., $\mathbf{j}_0$ contains those diagonal blocks of the Jordan form with diagonal elements zero. Then $\mathbf{j}_1$ has an $f$th root([5]), a matrix $\mathbf{j}$ such that $\mathbf{j}^f = \mathbf{j}_1$. Let $(\boldsymbol{\rho}_1, \boldsymbol{\rho}_0)$ be a partitioning of the columns of $\boldsymbol{\rho}$ conformal with that in $\mathrm{diag}\,(\mathbf{j}_1, \mathbf{j}_0)$. Then if $\mathbf{y}$ is the composite column vector

$$\mathbf{y} = [\boldsymbol{\rho}_1(\mathbf{j}\phi)^{f-1}, \mathbf{a}_1\boldsymbol{\rho}_1(\mathbf{j}\phi)^{f-2}, \cdots, \mathbf{a}_{f-1}\mathbf{a}_{f-2} \cdots \mathbf{a}_2\mathbf{a}_1\boldsymbol{\rho}_1(\mathbf{j}\phi)^0]'$$

where $\phi$ is an $f$th root of unity, one has

$$\mathbf{W}(h, \mathbf{A})y = P_f^{-1}\,\mathrm{diag}\,(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_f)\mathbf{y} = \mathbf{y}\mathbf{j}\phi.$$

Since $\phi$ may take any of the $f$ values $\exp\,[2\pi\, it/f]\,(t = 1, 2, \cdots, f)$ one may assemble $f$ composite columns $\mathbf{y}$, one for each $\phi$, into rectangular matrix $Y$.

The columns of $Y$ are linearly independent, as one may show.

If $\mathbf{a}$ is nonsingular then $Y$ is square and nonsingular and

$$Y^{-1}\,\mathbf{W}(h, \mathbf{A})Y = \mathrm{diag}\,(\mathbf{j}\phi_0, \mathbf{j}\phi_1, \cdots, \mathbf{j}\phi_{f-1})$$

where the various $\phi_i$ are the $f$th roots of unity in the order determined by the columns of $Y$. Finally, if $\boldsymbol{\rho}_i'$ is a matrix which transforms $\mathbf{j}\phi_i$ into Jordan form, then $YD$, where $D = \mathrm{diag}\,(\boldsymbol{\rho}_0', \boldsymbol{\rho}_1', \cdots, \boldsymbol{\rho}_{f-1}')$, transforms $\mathbf{W}(h, \mathbf{A})$ into Jordan form.

---

([5]) If $g(\lambda)$ is analytic at $\lambda_0$ then $g(\lambda_0 + h)$ may be written as a power series in $h$ with coefficients determined by $g$ and $\lambda_0$. If matrix $M = \lambda_0 I + H$ where $H$ is $H_s$ of (4.9) for some $s$ then $g(M)$ is given by the same series with $h$ replaced by $H$. If matrix $A$ is a direct sum of matrices of form $M$, as $\mathbf{j}_1$ is, then $g(A)$ is the direct sum of the separate summands $g(M)$. Since $H_s$ is nilpotent the various series terminate.

In discussing the case of singular $\mathbf{a}$ it is convenient to label as special vectors those composite vectors whose elements are simple vectors, all the simple vectors but one being zero. If $\mathbf{a}$ is singular, one or more of the $\mathbf{a}_i$ is singular. If $\mathbf{a}_k$ is singular and $z$ is a characteristic vector of $\mathbf{a}_k$ corresponding to root zero, then the special vector with $z$ as its $k$th element is characteristic for $\mathbf{W}(h, \mathbf{A})$ corresponding to root zero. Further every vector annihilated by $\mathbf{W}(h, \mathbf{A})$ may be written as a linear combination of such special vectors; for in order that $\mathbf{W}(h, \mathbf{A})\mathbf{x} = 0$, with $\mathbf{x}$ being some composite vector, it is necessary that the separate elements of $\mathbf{x}$ expanded into special vectors be annihilated. Indeed, vectors brought to zero by a power of $\mathbf{W}(h, \mathbf{A})$ may be written as linear combinations of special vectors, as one may establish by induction. Thus a set of base vectors which lead to the nilpotent part of the Jordan form for $\mathbf{W}(h, \mathbf{A})$ can be obtained from the simple vectors annihilated by $\mathbf{a}_k$ or $\mathbf{a}_k\mathbf{a}_{k-1}$ or $\cdots$ or $\mathbf{a}_k\mathbf{a}_{k-1}\cdots\mathbf{a}_{k-r}$ where $\mathbf{a}_k$ is singular, $0 \leqq r \leqq fm - 1$ and subscripts are considered to be reduced modulo $f$.

If appropriate vectors annihilated by a power of $\mathbf{W}(h, \mathbf{A})$ are assembled in proper order to a rectangular matrix Z, the matrix $(YD, Z)$ is a square nonsingular matrix which transforms $\mathbf{W}(h, \mathbf{A})$ into Jordan form.

7. **Composite zero circulants.** It is interesting and useful to exhibit what part of the solution of composite zero circulants can be performed explicitly. The basic relation is the analogue of (5.1),

$$\mathbf{A} = \mathbf{x}(n)\mathbf{r}, \qquad \mathbf{x}(n) = [I, I, \cdots, I]'. \tag{7.1}$$

This shows that any composite vector $\mathbf{y}$ is carried by $\mathbf{A}$ into a composite vector in the space $S$ spanned by the columns of $\mathbf{x}(n)$:

$$\mathbf{A}\mathbf{y} = \mathbf{x}(n)[\mathbf{r}\mathbf{y}].$$

Thus a composite vector whose elements are simple $m$th order column vectors either lies in $S$, is carried into $S$ by $\mathbf{A}$, or is a characteristic vector of $\mathbf{A}$ not in $S$ corresponding to proper value zero. If there are $k$ linearly independent characteristic vectors in this last class, let $\mathbf{Z}$ be a $k \times nm$ matrix whose columns span the space of these vectors.

If $N$ is any nonsingular $m$th order square matrix the columns of $\mathbf{x}(n)N$ span $S$. Finally let $\mathbf{Y}$ be an $[(n-1)m - k] \times nm$ matrix whose columns span the remaining space of vectors carried by $\mathbf{A}$ into a nonvanishing vector in $S$. Assemble these three rectangular matrices into nonsingular square matrix $R_0$.

$$R_0 = [\mathbf{x}(n)N, \mathbf{Y}, \mathbf{Z}]. \tag{7.2}$$

Then

$$\mathbf{A}R_0 = [\mathbf{x}(n)\mathbf{w}(n, \mathbf{A})N, \mathbf{x}(n)\mathbf{r}\mathbf{Y}, \mathbf{0}].$$

Put $N^{-1}\mathbf{w}(n, \mathbf{A})N = \mathbf{a}$ to obtain

$$\mathbf{A}R_0 = R_0 \begin{bmatrix} \mathbf{a} & N^{-1}\mathbf{rY} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

the elements of the composite third order matrix factor on the right being rectangular matrices of proper orders.

One may now specialize $N$ to be a matrix transforming $\mathbf{w}(n, \mathbf{A})$ to Jordan form: $\mathbf{a} = \operatorname{diag}[W_1, W_0]$ where $W_0$ is nilpotent and $W_1$ is nonsingular. If $B_1$ and $B_2$ give a conformal partitioning of $N^{-1}\mathbf{rY}$, then

$$R_0^{-1}\mathbf{A}R_0 = \begin{bmatrix} W_1 & \mathbf{0} & B_1 & \mathbf{0} \\ \mathbf{0} & W_0 & B_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

A further transformation using matrix $R_1$,

$$R_1 = \begin{bmatrix} I & \mathbf{0} & -W_1^{-1}B_1 & \mathbf{0} \\ \mathbf{0} & I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & I \end{bmatrix}, \tag{7.3}$$

eliminates $B_1$ to obtain

$$R_1^{-1}\ R_0^{-1}\ AR_0R_1 = \operatorname{diag}[W_1, M_1, \mathbf{0}]$$

where

$$M_1 = \begin{bmatrix} W_0 & B_2 \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{7.4}$$

The transformation of $\mathbf{A}$ to Jordan form is completed by solving $M_1$. Note that if $\mathbf{w}(n, \mathbf{A})$ is nonsingular, $M_1$ is vacuous and $\mathbf{A}$ has been solved.

If $R_2$ transforms $M_1$ to Jordan form $V_0$,

$$R_2^{-1}M_1R_2 = V_0 \tag{7.5}$$

then

$$R_3 = \operatorname{diag}[I, R_2, I]$$

transforms $\operatorname{diag}[W_1, M_1, 0]$ to Jordan form.

The only proper value of $M_1$ is zero. If $(y_1, y_2)'$ is a characteristic vector, $y_1 \neq 0$. For if $(0, y_2)$ were characteristic for $M_1$ then the appropriately expanded vector $(0, 0, y_2, 0, 0)'$ would be characteristic for $\mathbf{A}$ and would lie in the space spanned by the columns of $\mathbf{Z}$ and not those of $\mathbf{Y}$.

Hence $B_2 y_2 = 0$ has no nontrivial solutions, and $B_2$ has no more columns than rows. The dimension $k$ of the space spanned by $\mathbf{Y}$ is seen to be no greater than the dimension of $W_0$. Thus $k < m$ and matrix $M_1$ is of order no greater than $2m$. The solution of $nm$th order composite zero circulant $\mathbf{A}$ has been reduced to the solution of $m$th order matrix $\mathbf{w}(n, \mathbf{A})$ and matrix $M_1$ of order no more than $2m$.

Something further may be uncovered using the nilpotence of $W_0$. Simple induction establishes

$$M_1^p = \begin{bmatrix} W_0^p & W_0^{p-1}B_2 \\ 0 & 0 \end{bmatrix}$$

so that if $W_0^{p-1} = 0$ then $M_1^p = 0$. The order of the largest canonical diagonal block in Jordan matrix $W_0$ is therefore no greater than $(p - 1)$ and the largest block in the form for $M_1$ of order no greater than $p$.

Note that if $y, M_1 y, M_1^2 y, \cdots, M_1^q y$ span a canonical invariant subspace of $M_1$ then $M_1 y, M_1^2 y, \cdots, M_1^q y$ span an invariant subspace of $W_0$. Hence, there is a correspondence in which the individual blocks in the Jordan form for $M_1$ are of the same order or one order higher than the corresponding blocks of $W_0$.

The above results on simple and composite zero circulants are summarized in the following theorem:

THEOREM 7.1. *If the submatrix elements of nth order composite zero-circulant matrix* $\mathbf{A}$ *are mth order square matrices* $\mathbf{a}_{ij}$, *and if a Jordan form for* $\mathbf{w}(n, \mathbf{A}) = \mathbf{a}_{11} + \mathbf{a}_{12} + \cdots \mathbf{a}_{1n}$ *is written* $\operatorname{diag}[W_1, W_0]$ *where* $W_1$ *is nonsingular and* $W_0$ *is nilpotent, then a Jordan form for* $\mathbf{A}$ *is* $\operatorname{diag}[W_1, V_0, 0]$ *where* $V_0$ *is nilpotent, the order of* $V_0$ *is no more than twice the order of* $W_0$, *and there is a correspondence in which the individual diagonal blocks of* $V_0$ *are either of the same or one order higher than corresponding blocks of* $W_0$.

*A matrix transforming* $\mathbf{A}$ *to Jordan form is* $R_0 R_1 R_3$, *the matrices* $R_i$ *being defined in equations* (7.2), (7.3), *and* (7.5) *respectively.*

It is noteworthy that the statements in the theorem about the relations between $W_0$ and $V_0$ are precise. For one may readily construct examples of pairs of nilpotent matrices $W_0$ and $V_0$ with any of the permissible correspondences.

8. **Nonprime circulants.** There remains the case of $n$th order $g$-circulant matrices with $g \neq 0$ and $g$ having a factor greater than 1 in common with $n$. It is

possible to reduce this case to the case $g = 0$ by regrouping the submatrix elements of $A$ into (larger) submatrices. The theory of §7 is then applied.

Let $A$ be a (composite or ordinary) $n \times n$ $g$-circulant with $(g, n) > 1$. If $g = \gamma h$ and $n = \delta h$, then $\delta g = \gamma n$, and

$$P_n^{\delta} A = P_n^{\delta-1} A P_n^{g} = \cdots = A P_n^{\delta g} = A.$$

The submatrices needed are $\delta$ times as large as the elements of $A$ in the original partitioning (in which $A$ is a $g$-circulant), i.e., $m' = \delta m$.

9. **Generalizations.** An immediate generalization of the above theory is obtained by replacing $P$ in the discussion by any similar matrix $Q$, i.e., any nonderogatory $Q$ having the same roots as $P$. For then there exists a transforming matrix $S$ such that

$$Q = S^{-1}PS.$$

If $A$ has the $g$-circulant property with respect to $Q$, i.e., if

$$QA = AQ^{g},$$

then similar matrix $SAS^{-1}$ is truly $g$-circulant, i.e., $g$-circulant with respect to $P$: $P(SAS^{-1}) = (SAS^{-1})P^{g}$.

The theory of $g$-circulant matrices may be expected to remain valid in number fields other than the complex field provided that unity has $k$ distinct $k$th roots for every $k$, $2 \leqq k \leqq n$.

Since the results of this paper flow almost exclusively from the equation $PA = AP^{g}$, many generalizations suggest themselves at once.

10. **Applications to dynamical systems.** The determination of the normal modes and natural frequencies of oscillation of a lumped parameter electrical or mechanical system requires the calculation of the roots and characteristic vectors of an appropriate matrix. If the system is sufficiently symmetrical that matrix may well be circulant(6).

For definiteness consider a system of $n$ point masses $m_j$ $(j=1, \cdots, n)$ interconnected by springs and constrained to having but one degree of freedom each. If $q_j$ is a generalized coordinate locating mass $m_j$ so that the kinetic energy of the system is

$$\frac{1}{2} \sum_{j=1}^{n} m_j q_j^2$$

and if the potential energy stored in perfectly elastic, massless springs is represented by

(6) Similar applications are considered by Egerváry [1] and at some length by Rutherford [5]. See also Whyburn [7] for an application of circulants with variable elements.

$$\frac{1}{2}\sum_{j=1}^{n}\sum_{h=1}^{n} k_{jh}(q_j - q_h)^2 + \frac{1}{2}\sum_{j=1}^{n} k_{jj}q_j^2,$$

then the equation of motion for the $j$th mass is

$$m_j\ddot{q}_j + \sum_{h=1}^{n} k_{jh}(q_j - q_h) + k_{jj}q_j = 0. \tag{10.1}$$

Here nonnegative constants $k_{jh}$ are geometrically modified spring constants which by the Newtonian equality of action and reaction form a symmetric set,

$$k_{jh} = k_{hj}.$$

The masses are moving in a normal mode at natural frequency $\omega$ if

$$q_j = q_j^{(0)}e^{i\omega t}, \qquad j = 1, 2, \cdots, n,$$

with appropriate constants $q_j^{(0)}$. Here $i = \sqrt{-1}$. Substitution of these expressions for $q_j$ into equation (10.1) shows that frequency $\omega$ is the square root of a characteristic root and the $q_j^{(0)}$ components of the corresponding characteristic vector of the dynamical matrix

$$A = M^{-1}(K - S)$$

where

$$M = \operatorname{diag}(m_1, m_2, \cdots, m_n),$$
$$K = \operatorname{diag}(K_1, K_2, \cdots, K_n),$$
$$K_j = \sum_{h=1}^{n} k_{jh},$$

and

$$S = (k_{ij}) - \operatorname{diag}(k_{11}, k_{22}, \cdots, k_{nn}).$$

As a first example consider $n$ equal masses constrained to move on a circle, connected to one another and to fixed points by springs in a completely symmetrical way so that, in equilibrium, the masses are equally spaced around the circle and the system of springs and masses appears the same viewed from each mass. Small motions of the $j$th mass are then governed by an equation of the form of (10.1) with $q_j$ the displacement of that mass along the circle.

The system appearing the same from each mass means that

$$k_{j,j+h} = k_{j+1,j+h+1}$$

for every $h$ and $j$. Thus $S$ is 1-circulant and, since $K$ and $M$ are here scalar matrices, $A$ is 1-circulant.

If the constraint of the masses to the circle is removed, the equations of small three-dimensional motions have the same appearance as (10.1) with $q_j$ replaced by $\mathbf{q}_j = (x_j, y_j, z_j)'$, $m_j$ replaced by $m_jI_3$ and the $k_{jh}$ replaced by appropriate third-order matrices $\mathbf{k}_{jh}$. It then appears that $S$ and $K$ are composite 1-circulant and, since $M$ is again a scalar matrix, $A$ is composite 1-circulant also.

One may see that equal masses arranged symmetrically as though strung on a necklace spiralling around a torus also give rise to a composite 1-circulant dynamical matrix.

For a different example consider two parallel rows of equal and equally spaced masses, numbered down one row and back up the other, the masses each being connected by springs to masses in the opposite row. (Lateral oscillation of a truss might be approximated in this manner.) Away from the ends of the rows, the spring arrangement being the same as viewed from each mass means

$$k_{j,h} = k_{j+1,h-1}.$$

Appropriate connection of the end masses to fixed points permits this relation to hold for all $j$ and $h$. It follows that $S$ is $(-1)$-circulant. As before, $K$ and $M$ are scalar matrices so that the roots of dynamical matrix $A$ are simple functions of the roots of $S$ and the vectors of $A$ and $S$ agree.

It is of interest to determine which circulant matrices could arise from dynamical matrices. More specifically, if $S$ is symmetric and has zeros on its main diagonal, for what $g$ can $S$ be $g$-circulant? In answering this question (completely) we shall show that $S$ is a composite 1-circulant.

Write $s_{ij}$ for the $(i,j)$th element of $S$. The hypotheses are:

$$\begin{aligned} &\text{Symmetry:} && s_{ij} = s_{ji}, \\ &\text{Zero diagonal:} && s_{ii} = 0, \\ &g\text{-circulant:} && s_{ij} = s_{i-1,j-g} \quad (i,j = 1,\cdots,n). \end{aligned}$$

With a special notation for elements of the first row of $S$

$$s_{ij} = \sigma_j,$$

the $g$-circulant property implies

$$s_{ij} = \sigma_{j-g(i-1)}.$$

The zero diagonal property gives

$$\sigma_{g-i(g-1)} = 0.$$

If $(g-1)$ and $n$ are relatively prime, $(g-1)$ has an inverse modulo $n$ and every positive integer no greater than $n$ may be represented as

$$g - i(g-1) \mod n$$

for some $i$. Thus, if $S \neq 0$, $(g-1, n) = h \neq 1$.

In the matrix notation, since $S$ is $g$-circulant

$$SP^g = PS$$

and, for any $k$,

$$SP^{kg} = P^k S.$$

In particular, if $k = n/h$, $kg = [n(g-1)/h] + k$ so that

$$P^{kg} = P^k$$

and

$$SP^k = P^k S.$$

This last may be interpreted to mean that $S$ is a composite 1-circulant matrix of order $h$ with submatrix elements of order $k = n/h$.

As a composite 1-circulant $S$ is symmetric if its elements are symmetric and if, for any $i$,

$$\boldsymbol{\sigma}_{i+1} = \boldsymbol{\sigma}_{i-1},$$

i.e., pairs of elements of the first row at the same distance from $\boldsymbol{\sigma}_1$ are equal.

In summary, $n$th order nonvanishing matrix $S$ can be symmetric, have a zero main diagonal, and be $g$-circulant if $(g-1, n) = h \neq 1$. If so, $S$ is also composite 1-circulant of order $h$ with symmetric submatrix elements, pairs of which at the same distance from the main diagonal are equal.

As a final example consider twelve masses equally spaced around a circle and constrained to move along frictionless tracks which lie in the plane of the circle and normal to its circumference. Number the masses 1 through 12 clockwise, as in the figures.

Let springs of equal modified spring constant $k_{ij} = x$ connect the masses whose representative points are joined by straight lines in Figure 1. Springs with spring constant $y$ connect the masses joined by the lines of Figure 2, and similarly for springs with spring constant $z$ and $w$ in Figures 3 and 4, respectively.

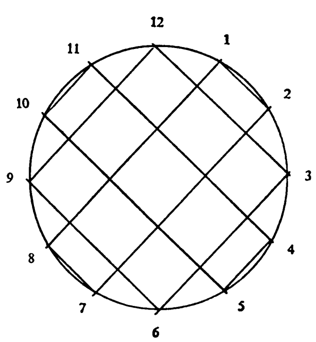


FIGURE 1
Springs with $k_{ij} = x$

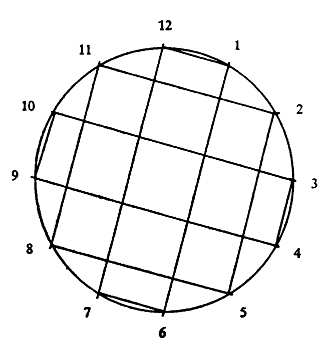


FIGURE 2
Springs with $k_{ij} = y$

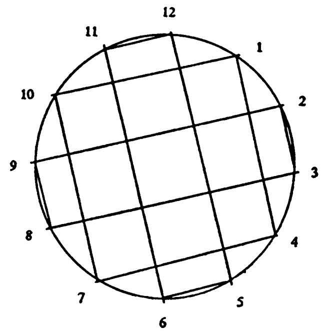


FIGURE 3
Springs with $k_{ij} = z$

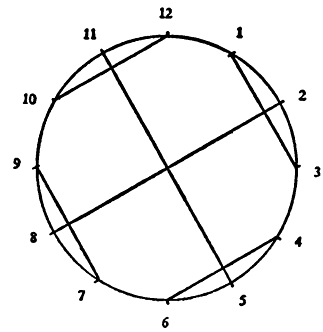


FIGURE 4
Springs with $k_{ij} = w$

These arrangements of springs look the same from each triple of masses (1, 2, 3), (4, 5, 6), (7, 8, 9), or (10, 11, 12). Their dynamical matrix $S$ is therefore composite 1-circulant with submatrix elements of order 3. The first composite row of $S$ reads

$$\begin{matrix} 0 & x & w & : & z & 0 & y & : & 0 & x & 0 & : & z & 0 & y \\ x & 0 & z & : & 0 & y & 0 & : & x & w & z & : & 0 & y & 0 \\ w & z & 0 & : & y & 0 & x & : & 0 & z & 0 & : & y & 0 & x \end{matrix} .$$

One may verify that as a noncomposite matrix $S$ is a 5-circulant of order 12. If the springs of constant $w$ are absent, if $w = 0$, $S$ is also $(-1)$-circulant.

## REFERENCES

1. E. Egerváry, *On hypermatrices whose blocks are commutable in pairs and their application in lattice dynamics*, Acta Sci. Math. (Szeged) **15** (1954), 211–222.

2. B. Friedman, *n-commutative matrices*, Math. Ann. **136** (1958), 343–347.

3. ———, *Eigenvalues of composite matrices*, Proc. Cambridge Philos. Soc. **57** (1961), 37–49.

4. F. A. Lewis, *Circulants and their groups*, Amer. Math. Monthly **67** (1960), 258–266.

5. D. E. Rutherford, *Some continuant determinants arising in physics and chemistry*, Proc. Roy. Soc. Edinburgh. Sect. A **62** (1951), 229–236; **63** (1951), 232–241.

6. E. Cesàro and G. Kowalewski, *Elementares Lehrbuch der algebraischen Analysis*, Teubner, Leipzig, 1904, pp. 25–26.

7. W. M. Whyburn, *A set of cyclically related functional equations*, Bull. Amer. Math. Soc. **36** (1930), 863–868.

STANFORD RESEARCH INSTITUTE,
MENLO PARK, CALIFORNIA