Displacement Operator Based Decompositions of Matrices Using Circulants or Other Group Matrices*

Paul D. Gader
Environmental Research Institute of Michigan
P.O. Box 8618
Ann Arbor, Michigan 48107

Submitted by Richard A. Brualdi

ABSTRACT

We show how an arbitrary square matrix can be expressed as sums of products of circulant and upper or lower triangular Toeplitz matrices, and as sums of products of matrices derived from finite groups (group matrices) and matrices which are "close" to group matrices. The results obtained are interesting from several points of view: they lead to different methods for computing linear transforms using FFTs or fast convolution algorithms, to faster methods for solving Toeplitz systems, to potential methods for mapping linear transforms to parallel computer architectures with interconnection scheme given by the group graph of a finite group, and possibly to matrix theoretic methods for expressing relationship between finite groups.

INTRODUCTION

In this paper, we show how an arbitrary square matrix can be expressed as sums of products of circulant and upper or lower triangular Toeplitz matrices, and as sums of products of matrices derived from finite groups (group matrices) and matrices which are "close" to group matrices. The results obtained are interesting from several points of view: they lead to different methods for computing linear transforms using FFTs or fast convolution algorithms, potential methods for mapping linear transforms to parallel

*This research was performed while the author was with the University of Wisconsin—Oshkosh.

655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

computer architectures with interconnection scheme given by the group graph of a finite group, and possible matrix theoretic methods for expressing relationships between finite groups.

Background

We begin with some definitions, conventions, and background. Let R be a ring with identity element 1, and let $M_n(R)$ denote the ring of $n \times n$ matrices over R. We consider matrices over rings so that our results will hold for matrices whose elements are matrices. We shall use the term linear operator for functions $f: M_n(R) \to M_n(R)$ with the property that $f(rA + sB) = rf(A) + sf(B)$, although module homomorphism is technically correct. We number the rows and columns of $n \times n$ matrices from 0 to $n-1$, and addition of subscripts will always be mod n. By the permutation matrix Q representing the permutation σ of $\{0, 1, \ldots, n-1\}$ we shall mean the $n \times n$ matrix

$$Q = (q_{ij}), \qquad \text{where} \quad q_{ij} = \begin{cases} 1 & \text{if} \quad \sigma(i) = j, \\ 0 & \text{otherwise.} \end{cases}$$

Circulant matrices can be used to model (periodic) time invariant digital filters or convolution operators. Furthermore, all circulant matrices over the complex numbers are simultaneously diagonalizable by the so-called Fourier matrices. Hence, linear transformations represented by circulant matrices can be computed by fast convolution or FFT algorithms. Expressing arbitrary matrices in terms of circulants can lead to methods for computing linear transforms using FFTs or fast convolution algorithms. The standard reference on circulants is [1], the books [2, 3] contain a good deal of information on fast convolutions and FFTs, and the papers [4–6] contain information concerning relations between Fourier and circulant matrices and their relationships to FFTs.

We now define group matrices. Let G be a finite group with $|G| = n$, and let $G = \{g_0, g_1, \ldots, g_{n-1}\}$ be any ordering of G, with g_0 denoting the identity element. Since the group elements will be used in subscripts, we identify G with the set $\{0, 1, \ldots, n-1\}$ in the obvious way. We let jk denote the element $g_j g_k$, so if $g_j g_k = g_m$, then we write $jk = m$. We also write j^s for g_j^s. This convention keeps the appearance of our results cleaner, but care must be taken, since, for example, $j(k+m) \neq jk + jm$ in general, and the identity is denoted by 0, i.e. $jj^{-1} = 0$. A group matrix (for G) over R is an $n \times n$ matrix $A = (a_{ij})$ with the property that $a_{i,j} = a_{ki,kj}$ for every $k \in G$.

A group matrix is defined relative to some group. Throughout this paper, when we refer to a group matrix, unless indicated otherwise, we shall be considering an arbitrary but fixed group G, so no explicit reference to G will

be made. We denote the set of group matrices by $R[G]$. The set $R[G]$ is closed under the operations of scalar multiplication (scalars coming from R), matrix addition, and matrix multiplication. If R is the complex numbers, then $R[G]$ is a regular representation of the group algebra of G over the complex numbers. Several authors have studied properties of group matrices; see e.g. [7–11].

An $n \times n$ circulant matrix is a group matrix for the cyclic group of order n. If

$$D_3 = \langle r, f \mid r^3 = f^2 = e,\ rfrf = e \rangle = \{e, r, r^2, f, fr, fr^2\}$$

denotes the dihedral group, then a group matrix for D_3 (relative to the above ordering) has the form

$$\left[\begin{array}{ccc:ccc} 0 & 1 & 2 & 3 & 4 & 5 \\ 2 & 0 & 1 & 4 & 5 & 3 \\ 1 & 2 & 0 & 5 & 3 & 4 \\ \hdashline 3 & 4 & 5 & 0 & 1 & 2 \\ 4 & 5 & 3 & 2 & 0 & 1 \\ 5 & 3 & 4 & 1 & 2 & 0 \end{array}\right],$$

where only the subscripts have been indicated for clarity. Note that, as is the case for circulants, group matrices are completely determined by row 0.

Group matrices represent linear transformations which are translation invariant with respect to group graphs, or Cayley networks. This is analogous to the (periodic) time invariance property of circulants. If Δ is a generating set for the group G, then Cayley network of G is a directed graph, $D_\Delta(G)$, with vertex set G having the property that (g, h) is an arc if and only if there is a $\delta \in \Delta$ such that $g = \delta h$ [12]. If G is the cyclic group of integers under addition mod n and $\Delta = \{1, -1 \equiv n-1\}$, then $D_\Delta(G)$ is the "circle" graph. If $G = D_3$ and $\Delta = \{r, f\}$, then $D_\Delta(G)$ is the directed graph shown in Figure 1. If we think of column vectors of length n as being functions defined on the vertex set of $D_\Delta(G)$, then if A is a group matrix, $x \to Ax$ "looks" the same at every point. For example, if $A = \text{circ}(1, 1, 0, \ldots, 0, 1)$ and $y = Ax$, then $y_i = x_{i-1} + x_i + x_{i+1}$ regardless of i. This is cyclic convolution. If A is the group matrix for D_3 with first row $(1, 0, 1, 1, 0, 0)$ and $y = Ax$, then $y_g = x_{gr^2} + x_g + x_{gf}$ for every $g \in D_3$. This has been called group convolution [13, 14] and, if R is the complex numbers, is multiplication in the group algebra of G over R. Several authors have proposed using Cayley networks of both abelian and nonabelian groups as models for interconnection schemes for parallel computer architectures [15–18]. Since group matrices represent highly regular transformations on these structures, the decompositions devel-

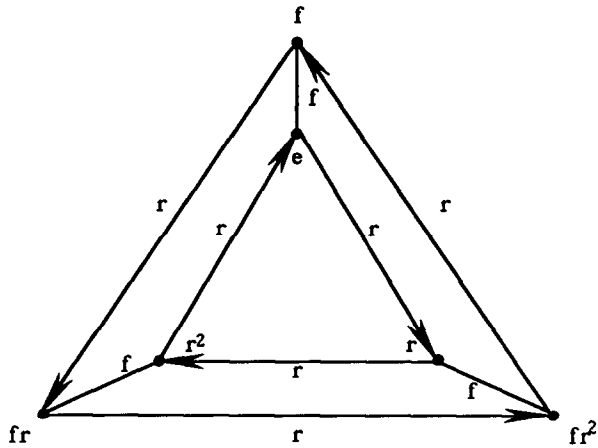

FIG. 1. Cayley network for dihedral group D_3.

oped in this paper could yield efficient methods for computing linear transforms on such architectures. Others have demonstrated how group convolutions can be useful for feature detection in image processing applications [19].

Preview

We now preview the results developed in this paper somewhat more precisely and discuss relationships to previous work. As mentioned in the first paragraph, we will develop methods for expressing arbitrary square matrices over R as either sums of products of circulant matrices and upper or lower triangular Toeplitz matrices, or sums of products of group matrices and matrices close to group matrices. That is, given a finite group G with $|G| = n$ and an $n \times n$ matrix A over R, we show how one can easily construct group matrices [using the singular value decomposition (SVD) if R is the real or complex numbers] $Y_1, Y_2, \ldots, Y_\alpha, K$, and matrices "close" to group matrices $X_1, X_2, \ldots, X_\alpha$ such that

$$A = \sum_{m=1}^{\alpha} X_m Y_m + K. \tag{1}$$

Notice that since $R[G]$ is closed under addition and multiplication, there is no hope of obtaining an expression of the form (1) with $X_1, \ldots, X_\alpha$ also group matrices unless $A \in R[G]$, in which case we shall obtain $A = K$.

As we have observed, circulants are special cases of group matrices. Thus, we could develop expressions of the form (1) for arbitrary group matrices and then cite the circulants as a special case. We prefer to go from the special to the general, that is, we first develop expressions in terms of circulants and then in terms of group matrices for an arbitrary finite group. We choose this approach for two reasons: one is that circulants arise frequently enough to warrant special attention, and the other is that the development is easier to understand for circulants than it is for arbitrary group matrices, so considering circulants first makes the exposition clearer.

Our work represents generalizations of work done by Kailath et al. [20, 21], who showed how to write an arbitrary square matrix over a ring as sums of products of upper and lower triangular matrices. The authors and others used these results to obtain several inversion formulas for matrices "close" to Toeplitz matrices using generalizations of the Levinson recursions; see e.g. [22, 23]. It was suggested in [20, 21, 24, 25] that these Toeplitz expressions could be used to compute arbitrary linear transforms using convolutions or FFTs. It was also suggested, but not carried out, that similar expressions could be developed for circulants [20]. We feel that the circulant expressions we have developed may be more directly implemented using fast convolution algorithms because most of these algorithms are based on circular convolutions, represented by circulants, rather than linear convolutions, represented by Toeplitz matrices [2]. In fact, we have used these expressions to derive an analogue of the Gohberg-Semencul formula for the inverse of a Toeplitz matrix. This new formula yields computational savings of 35% in the second phase of a superfast Toeplitz solver [27, 28]. We are not aware of any work in this area related to arbitrary group matrices.

Apart from the different potential applications resulting from our work, several differences arise when attempting to apply previous methods to circulants and other group matrices. To see what they are, we briefly review what was done in [20]. A linear "displacement" operator, $^{\downarrow}$, is defined on $M_n(R)$ to produce a displacement matrix $^{\downarrow}A$ using

$$^{\downarrow}A = A - ZAZ^t,$$

where Z is the lower shift matrix

$$Z := \begin{bmatrix} 0 & & & & & \\ 1 & 0 & & & & \\ & 1 & 0 & & & \\ & & \ddots & \ddots & & \\ & & & 1 & 0 \end{bmatrix}.$$

If T is Toeplitz, then $\operatorname{rank}(\lrcorner T) \leqslant 2$. It is shown that if $x = (x_0, x_1, \ldots, x_{n-1})^t$ and $y = (y_0, y_1, \ldots, y_{n-1})^t$, then we can write $\lrcorner A$ as the outer product $\lrcorner A = xy^t$ if and only if $A = L(x)U(y^t)$, where $L(x)$ is the $n \times n$ lower triangular Toeplitz matrix with first column equal to y and $U(y^t) = L(y)^t$. Hence, if

$$\lrcorner A = \sum_{m=1}^{\alpha} x_m y_m^t,$$

then

$$A = \sum_{m=1}^{\alpha} L(x_m)U(y_m^t)$$

is the desired expression. The main differences that arise are:

(1) Since $R[G]$ is closed under addition and multiplication, there is no hope of obtaining similar expressions entirely in terms of group matrices. It is not clear what form the expressions should take for an arbitrary group G.

(2) The "natural" displacement operators for group matrices have nontrivial kernels, which is not true of the operator $\lrcorner$. Thus, there are matrices which are not displacement matrices, and furthermore we no longer have a property analogous to "if $A = L(x)U(y^t)$, then $\lrcorner A = xy^t$" for arbitrary x and y. Moreover, the form of the displacement operator must be modified for group matrices corresponding to noncyclic groups.

(3) Certain orthogonality relations must hold between the column vectors x and y for the outer product xy^t to represent a group displacement matrix, or more generally, for a sum of outer products to represent a group displacement matrix. These orthogonality relations are essential ingredients of the expressions developed in this paper, but are not required at all in the Toeplitz case.

The statements made in (2) and (3) are made precise and thoroughly investigated in the following sections.

The rest of the paper is organized as follows: We first define displacement operators and develop orthogonality relations for circulants, and then develop expressions for matrices in terms of circulants. We then parallel the development for arbitrary group matrices, pointing out the differences and similarities as we go. We conclude the paper with a summary and some suggestions for further research.

THE CIRCULANT CASE

Definitions

Throughout this section we let $P := \mathrm{circ}(0,1,0,\ldots,0)$. Define the linear operators $\mathscr{E}$ and $\tilde{\mathscr{E}}$ on the algebra of all $n \times n$ matrices by

$$\mathscr{E}(A) := A - P^t A P \quad \text{and} \quad \tilde{\mathscr{E}}(A) := A - PAP^t.$$

Note that the i,j element of $\mathscr{E}(A)$ is $a_{ij} - a_{i-1,j-1}$ and that of $\tilde{\mathscr{E}}(A)$ is $a_{ij} - a_{i+1,j+1}$. Furthermore, $\mathscr{E}(A) = 0$ and $\tilde{\mathscr{E}}(a) = 0$ are both equivalent to A being circulant, that is, $\ker \mathscr{E} = \ker \tilde{\mathscr{E}} = \mathscr{C}_n$, the set of circulants. Let S be the reversal matrix, i.e. the matrix $[e_n, e_{n-1}, \ldots, e_1]$, and let $x = (x_0, x_1, \ldots, x_{n-1})^t$, $y = (y_0, y_1, \ldots, y_{n-1})^t$, $\tilde{x} = Sx$, and $\tilde{y} = Sy$. We first derive the aforementioned orthogonality relations.

LEMMA 1. *If* $\mathscr{E}(A) = xy^t$, *then* $(P^k x)^t y = 0$ *for* $k = 0,1,\ldots,n-1$. *More generally, if* $\mathscr{E}(A) = \sum_{m=1}^{\alpha} x_m y_m^t$, *then* $\sum_{m=1}^{\alpha} (P^k x_m)^t y_m = 0$ *for* $k = 0,1,\ldots,n-1$.

Proof. We first consider the case $\mathscr{E}(A) = xy^t$. Setting the diagonal elements of $\mathscr{E}(A)$ and xy^t equal yields

$$\begin{aligned} a_{00} - a_{n-1,n-1} &= x_0 y_0, \\ a_{11} - a_{00} &= x_1 y_1, \\ &\vdots \\ a_{n-1,n-1} - a_{n-2,n-2} &= x_{n-1} y_{n-1}. \end{aligned}$$

Since the left hand sides sum to zero, the right hand sides must also; thus $x^t y = 0$. Similarly, we can set the *circulant subdiagonals* equal to obtain $(P^k x)^t y = 0$ for $k > 0$. For example, if a matrix B is 4×4, then the second circulant subdiagonal is $\{b_{20}, b_{31}, b_{02}, b_{13}\}$. For k between 1 and $n-1$ this yields

$$\begin{aligned} a_{k,0} - a_{k-1,n-1} &= x_k y_0, \\ a_{k+1,1} - a_{k,0} &= x_{k+1} y_1, \\ &\vdots \\ a_{k+n-1,n-1} - a_{k+n-2,n-2} &= x_{k+n-1} y_{n-1}, \end{aligned}$$

which imply that $(P^k x)^t y = 0$, since $P^k x = (x_k, x_{k+1}, \ldots, x_{k+n-1})^t$.

The proof for the more general case, $\alpha > 1$, is essentially the same. Letting $x_m := (x_0^{(m)}, x_1^{(m)}, \ldots, x_{n-1}^{(m)})$ we obtain, for $k = 0, 1, \ldots, n-1$,

$$a_{k,0} - a_{k-1,n-1} = x_k^{(1)} y_0^{(1)} + x_k^{(2)} y_0^{(2)} + \cdots + x_k^{(\alpha)} y_0^{(\alpha)},$$

$$a_{k+1,1} - a_{k,0} = x_{k+1}^{(1)} y_1^{(1)} + x_{k+1}^{(2)} y_1^{(2)} + \cdots + x_{k+1}^{(\alpha)} y_1^{(\alpha)},$$

$$\vdots$$

$$a_{k+n-1,n-1} - a_{k+n-2,n-2} = x_{k+n-1}^{(1)} y_{n-1}^{(1)} + x_{k+n-1}^{(2)} y_{n-1}^{(2)} + \cdots + x_{k+n-1}^{(\alpha)} y_{n-1}^{(\alpha)}.$$

Summing the equations yields the result (since $k + n - 1 \equiv k - 1$). ■

LEMMA 2. *If* $\mathscr{E}(A) = \sum_{m=1}^{\alpha} \tilde{x}_m \tilde{y}_m^t$, *then* $\sum_{m=1}^{\alpha} (P^k x_m)^t y_m = 0$ *for* $k = 0, 1, \ldots, n-1$.

The proof of this lemma can be formulated using the proof of Lemma 1 as a guide.

Decomposition Theorems

We now state the decomposition theorems and corollaries and then give the proofs.

THEOREM 3. *If* $\mathscr{E}(A) = xy^t$, *then there exist circulant matrices* Y *and* C_l, *and a lower triangular Toeplitz matrix* $L(x)$, *such that*

$$A = L(x)Y + C_l.$$

In this case we have

$$Y := \operatorname{circ}(y^t) = \operatorname{circ}(y_0, y_1, \ldots, y_{n-1})$$

and

$$C_l = \operatorname{circ}(a_{n-1,1}, a_{n-1,2}, \ldots, a_{n-1,n-1}, a_{n-1,0}).$$

COROLLARY 4. *If* $\mathscr{E}(A) = \sum_{m=1}^{\alpha} x_m y_m^t$, *then*

$$A = \left[\sum_{m=1}^{\alpha} L(x_m) Y_m \right] + C_l,$$

where the Y_m *and* C_l *are defined as in Theorem* 3.

THEOREM 5. *If* $\tilde{\mathscr{E}}(A) = \tilde{x}\tilde{y}^t$, *then there exist circulant matrices* Y *and* C_u *such that*

$$A = U(x^t)Y^t + C_u,$$

where

$$Y = \operatorname{circ}(y^t)$$

and

$$C_u = \operatorname{circ}(a_{00}, a_{01}, \ldots, a_{0,n-1}).$$

COROLLARY 6. *If* $\tilde{\mathscr{E}}(A) = \sum_{m=1}^{\alpha} \tilde{x}_m \tilde{y}_m^t$, *then*

$$A = \left[\sum_{m=1}^{\alpha} U(x_m^t) Y_m^t \right] + C_u,$$

where Y_m *and* C_u *are defined as in Theorem* 5.

In the case of matrices over the real or complex numbers, both decompositions can be obtained directly from the SVD of the displacement matrix. The decomposition given in Corollary 6 is more pleasing esthetically, since the lone circulant term, C_u, is expressed in terms of the first row of the original matrix A rather than the last.

We now prove Theorem 3 and Corollary 4. The proofs of Theorem 5 and Corollary 6 are similar.

Proof of Theorem 3 and Corollary 4. If $\mathscr{E}(A) = xy^t$, then the orthogonality relations of Lemma 1 hold. The i, j element of $L(x)Y$ is $\sum_{k=0}^{i} x_{i-k} y_{j-k}$.

Hence, if $i \neq 0$, the i,j element of $\mathscr{C}(L(x)Y)$ is

$$\sum_{k=0}^{i} x_{i-k}y_{j-k} - \sum_{k=0}^{i-1} x_{i-1-k}y_{j-1-k} = x_i y_j,$$

which is precisely the i,j element of xy^t. If $i = 0$, then the i,j element of $\mathscr{C}(L(x)Y)$ is

$$x_0 y_j - \left(\sum_{k=0}^{n-1} x_{n-1-k}y_{j-1-k} \right).$$

This term will equal $x_0 y_j$ for every $j = 0, 1, \ldots, n-1$ if and only if

$$\sum_{k=0}^{n-1} x_{n-1-k}y_{j-1-k} = \sum_{k=0}^{n-1} x_{k-j}y_k = 0, \qquad j = 0, 1, \ldots, n-1,$$

which is true by Lemma 1. Hence $\mathscr{C}(L(x)Y) = xy^t = \mathscr{C}(A)$, so there is a circulant matrix C_l such that $A = L(x)Y + C_l$ (since $\ker \mathscr{C} = \mathscr{C}_n$). Furthermore, the last row of $L(x)Y$ is

$$\left[(P^{n-1}x)^t y, (P^{n-2}x)^t y, \ldots, (Px)^t y, x^t y \right],$$

which is the zero row, again using the orthogonality relations of Lemma 1. Thus A and C_u must have the same last row, which implies that C_u has the form given in the statement of the theorem.

The proof of the corollary is essentially the same. Briefly, the i,j element of $\Sigma_{m=1}^{\alpha} L(x_m)Y_m$ is $\Sigma_{m=1}^{\alpha} \Sigma_{k=0}^{i-1} x_{i-k}^{(m)} y_{j-k}^{(m)}$, and so the i,j element of $\mathscr{C}[\Sigma_{m=1}^{\alpha} L(x_m)Y]$ is $\Sigma_{m=1}^{\alpha} x_i^{(m)} y_j^{(m)}$ if $i \neq 0$ and

$$\sum_{m=1}^{\alpha} x_0^{(m)} y_j^{(m)} - \sum_{m=1}^{\alpha} \left[\sum_{k=0}^{n-1} x_{n-1-k}^{(m)} y_{j-1-k}^{(m)} \right]$$

if $i = 0$. Equality of elements of $\Sigma_{m=1}^{\alpha} x_m y_m$ and $\mathscr{C}[\Sigma_{m=1}^{\alpha} L(x_m)Y_m]$ will hold if and only if

$$\sum_{m=1}^{\alpha} \left[\sum_{k=0}^{n-1} x_{n-1-k}^{(m)} y_{j-1-k}^{(m)} \right] = 0 \qquad \text{for} \quad j = 0, 1, \ldots, n-1,$$

or

$$\sum_{m=1}^{\alpha}\left[\sum_{k=0}^{n-1} x_{k-j}^{(m)} y_k^{(m)}\right] = 0,$$

which is true by Lemma 1. ∎

In contrast to the Toeplitz case, it is not generally true that if $A = L(x)Y + C$ then $\mathscr{E}(A) = xy^t$, because the orthogonality relations do not hold in general. When they do hold, we have the following partial converses, the proofs of which are essentially contained in the above arguments.

THEOREM 7. *If* $\{x_m\}_{m=1}^{\alpha}$ *and* $\{y_m\}_{m=1}^{\alpha}$ *are families of n-vectors satisfying* $\sum_{m=1}^{\alpha}(P^k x_m)^t y_m = 0$ *for* $k = 0, 1, \ldots, n-1$, *then* $A = \sum_{m=1}^{\alpha} L(x_m)Y_m + C$ *implies* $\mathscr{E}(A) = \sum_{m=1}^{\alpha} x_m y_m^t$, *and* $A = \sum_{m=1}^{\alpha} U(x_m^t)Y_m^t + C$ *implies* $\mathscr{E}(A) = \sum_{m=1}^{\alpha} \tilde{x}_m \tilde{y}_m^t$.

We can combine Theorems 3, 5, and 7 to obtain a characterization of square matrices over $\mathbb{R}$ or $\mathbb{C}$ which are circulant displacement matrices.

THEOREM 8 (Characterization of circulant displacement matrices). *Let* $B \in M_n(\mathbb{C})$; *let* $P = \operatorname{circ}(0, 1, 0, \ldots, 0)$. *The following are equivalent*:

(1) *there exists a matrix A such that* $\mathscr{E}(A) = B$;
(2) *there exists a matrix* $\tilde{A}$ *such that* $\tilde{\mathscr{E}}(\tilde{A}) = B$;
(3) *there is an outer product representation*

$$B = \sum_{m=1}^{\alpha} u_m v_m^t$$

such that the relations $\sum_{m=1}^{\alpha}(P^k u_m)^t v_m = 0$ *hold for every* $k = 0, 1, \ldots, n-1$.

ARBITRARY GROUP MATRICES

We now turn our attention to the more general case of arbitrary group matrices. Of the several possible variations of decompositions, we consider the decompositions analogous to those obtained using the operator $\tilde{\mathscr{E}}$ in the previous section. We assume throughout that, unless otherwise indicated, G is an arbitrary but fixed finite group with $|G| = n$ and that $|R| \geqslant n$.

LEMMA 9. *If G is not a cyclic group, then there is no permutation matrix Q such that $R[G]$ is the kernel of the linear operator $\mathcal{S}$ defined by $\mathcal{S}(A) := A - QAQ^t$.*

Proof. We first show that if such a Q were to exist, then it would have to represent a certain type of permutation, called a color preserving automorphism. We then show that if Q represents any color preserving automorphism, then there exists a matrix $B \notin R[G]$ such that $B - QBQ^t = 0$.

Assume that Q is an $n \times n$ permutation matrix and that $R[G] \subset \ker(\mathcal{S})$. Let $A \in R[G]$ with n distinct elements of R in row 0. If Q represents the permutation σ, then $\sigma(0) = g0 = g$ for some $g \in G$. Since $A \in R[G]$, for every $j = 0, 1, \ldots, n-1$ we have $a_{0,j} = a_{g,gj}$. Since $A \in \ker(\mathcal{S})$, $a_{0,j} = a_{\sigma(0),\sigma(j)} = a_{g,\sigma(j)}$. By construction, $a_{g,k} = a_{g,m}$ if and only if $k = m$, which implies that $\sigma(j) = gj$ for every $j \in G$.

Since G is not cyclic, the set

$$C(g) = \{g^k \mid k = 0, 1, \ldots, n-1\}$$

is a proper subset of G. Let $m \in G \setminus C(g)$, and let s be the order of g, that is, $g^{s-1} \neq 0$ and $g^s = 0$. We now construct a matrix $B \in \ker(\mathcal{S}) \setminus R[G]$. Let $c \in R$, and take the first row of B to be $(c, c, \ldots, c)$. Define the elements in rows $g, g^2, \ldots, g^{s-1}$ by

$$b_{g^k, g^k j} := c, \qquad k = 0, 1, \ldots, s-1 \text{ and } j = 0, 1, \ldots, n-1.$$

Let $d \in R$ with $d \neq c$, and define the elements in rows $m, gm, g^2m, \ldots, g^{s-1}m$ by

$$b_{g^k m, g^k mj} = d, \qquad k = 1, 2, \ldots, s-1 \text{ and } j = 0, 1, \ldots, n-1.$$

The elements in the remaining rows (if there are any) can all be set equal to any constant $e \in R$. Since B is not completely determined by row 0, $B \notin R[G]$.

Let b_{ij} be any element of B. If $i = g^k$ for some k, then

$$b_{ij} = b_{g^k, j} = c = b_{g^{k+1}, gj} = b_{\sigma(i),\sigma(j)}.$$

Similarly, if $i = g^k m$ for some k, then

$$b_{ij} = b_{g^k m, j} = d = b_{g^{k+1} m, gj} = b_{\sigma(i),\sigma(j)}.$$

Finally, if $i \neq g^k$ and $i \neq g^k m$ for any k, then the same is true for gi, so $b_{ij} = b_{gi,gj} = b_{\sigma(i),\sigma(j)}$. Thus, in any case $b_{ij} = b_{\sigma(i),\sigma(j)}$, so $\mathscr{G}(B) = 0$. ■

The lemma implies that block circulants with circulant blocks cannot always be realized as $\ker \mathscr{G}$ for some $\mathscr{G}$ of the form $\mathscr{G}(A) = A - QAQ^t$, since they are group matrices of direct sums of cyclic groups, which are not always cyclic. Since we are constructing our decompositions over a ring R, we can, for example, consider block circulants over the complex numbers to be circulant matrices over the ring of $n \times n$ matrices over $\mathbb{C}$.

We now use the fact that each row of a group matrix is a permutation of the other rows to define a displacement operator with kernel $R[G]$. For each $i = 0, 1, \ldots, n-1$ let $k_i \in G$ such that $k_i i = i - 1 \pmod{n}$. We shall use the notation k_i on several occasions. Define the permutation τ_i by $\tau_i(j) = k_i^{-1} j$, and let Q_i be the permutation matrix corresponding to τ_i. We shall see in the proof of the next lemma that if $A \in R[G]$ and if r_i^t is the ith row of A, then $Q_i r_i = r_{i-1}$. Define the permutation operator $\mathscr{Q}$ by

$$\mathscr{Q}(A) := [Q_0 r_0, Q_1 r_1, \ldots, Q_{n-1} r_{n-1}]^t.$$

Define the linear operator $\mathscr{G}$ on $M_n(R)$ by $\mathscr{G}(A) := A - P\mathscr{Q}(A)$.

LEMMA 10. $\ker \mathscr{G} = R[G]$.

Proof. Let $A \in M_n(R)$. If r_i^t is the ith row of A, then

$$Q_i r_i = (a_{i,k_j^{-1}0}, a_{i,k_j^{-1}1}, \ldots, a_{i,k_i^{-1}(n-1)})^t.$$

If $A \in R[G]$, then $a_{i,k_i^{-1}j} = a_{i-1,j}$, so $Q_i r_i = r_{i-1}$. Hence,

$$A - P\mathscr{Q}(A) = [r_0, r_1, \ldots, r_{n-1}]^t - P[Q_0 r_0, Q_1 r_1, \ldots, Q_{n-1} r_{n-1}]^t = 0.$$

Conversely, assume $A \in \ker \mathscr{G}$. Since $i + 1 = k_{i+1}^{-1}$, the i, j element of $\mathscr{G}(A)$ is

$$0 = a_{i,j} - a_{i+1,k_{i+1}^{-1}j} = a_{i,j} - a_{k_{i+1}^{-1}i,k_{i+1}^{-1}j}.$$

Let $g \in G$. We show that $a_{i,j} = a_{gi,gj}$ for every i, j. Fix i and j, and note that we may assume that $gi = i + p$, where $0 < p < n - i$, since, if not, we

can show $a_{r,s} = a_{g^{-1}r,g^{-1}s}$, where $r = gi$ and $s = gj$. For each $m = 1,2,\ldots,p$ let

$$g(m) := k_{i+m}^{-1} k_{i+m-1}^{-1} \cdots k_{i+1}^{-1}.$$

Observe that $k_{i+m}^{-1} g(m-1) = g(m)$ and also that $gi = i + p$, which implies $k_{i+1}k_{i+2} \cdots k_{i+p} gi = i$, so $g = g(p)$. Hence,

$$a_{ij} = a_{g(1)i,g(1)j} = a_{g(2)i,g(2)j} = \cdots = a_{g(p)i,g(p)j} = a_{gi,gj}. \quad \blacksquare$$

Orthogonality Relations

We now show that orthogonality relations hold for arbitrary group matrices. For each $j = 0,1,\ldots,n-1$ let F_j be the $n \times n$ permutation matrix representing the permutation $\sigma_j(m) := mj$. Let $x, y \in R^n$ be column vectors.

LEMMA 11. *If* $\mathscr{G}(A) = \sum_{m=1}^{\alpha} x_m y_m^t$, *then* $\sum_{m=1}^{\alpha} x_m^t (F_j y_m) = 0$ *for* $j = 0,1,\ldots,n-1$.

Proof. Assume that $\mathscr{G}(A) = xy^t$. Let $j \in G$ be arbitrary but fixed. Recall that if $i \in G$, then k_i denotes the element of G with the property that $k_i i = i - 1$. For every $m \in G$ let $k(m) := k_m^{-1} k_{m-1}^{-1} \cdots k_1^{-1}$. Notice that $k(m) = m$, since $k(m)^{-1} m = k_1 k_2 \cdots k_m m = 0$. Setting corresponding entries of $\mathscr{G}(A)$ and xy^t equal as in Lemma 1 yields

$$\begin{aligned}
a_{0,j} - a_{1,k_1^{-1}j} &= x_0 y_j, \\
a_{1,k_1^{-1}j} - a_{2,k(2)j} &= x_1 y_{k(1)j} = x_1 y_{1j}, \\
a_{2,k(2)j} - a_{3,k(3)j} &= x_2 y_{k(2)j} = x_2 y_{2j}, \\
&\vdots \\
a_{n-1,k(n-1)j} - a_{0,j} &= x_{n-1} y_{k(n-1)j} = x_{n-1} y_{(n-1)j}.
\end{aligned}$$

Since the left hand sides sum to zero, the right hand sides do too. Furthermore, $F_j y = (y_j, y_{1j}, \ldots, y_{(n-1)j})^t$ shows that $x^t(F_j y) = 0$. The proof of the more general case can be constructed from this proof just as the proof of the general case of Lemma 1 was. $\blacksquare$

Group Matrix Decomposition

We now develop the decomposition of arbitrary matrices into sums of products of group matrices and matrices which are "close" to group matrices. We first describe these matrices and then show how they yield our decompositions.

Let $x, y \in R^n$ be column vectors. For every $i \in G$, let E_i be the $n \times n$ permutation matrix representing $\pi_i(j) := j^{-1}i$.

DEFINITION 12. Let $U_G(x)$ be the $n \times n$ matrix defined as follows:

(1) Row 0 of $U_G(x)$ is $(E_0 x)^t$.

(2) Row 1 of $U_G(x)$ is the same as $(E_i x)^t$ except that $x_0, x_1, \ldots, x_{i-1}$ are replaced by zero.

EXAMPLE. If G is the dihedral group D_3, then, using the same ordering of D_3 as before, $U_G(x)$ has the form

$$\left[\begin{array}{ccc|ccc} x_0 & x_2 & x_1 & x_3 & x_4 & x_5 \\ x_1 & 0 & x_2 & x_4 & x_5 & x_3 \\ x_2 & 0 & 0 & x_5 & x_3 & x_4 \\ \hline x_3 & x_4 & x_5 & 0 & 0 & 0 \\ x_4 & x_5 & 0 & 0 & 0 & 0 \\ x_5 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

In the circulant case, $U_G(x) = U(\tilde{x}^t)$.

In what follows, we shall use $\Gamma_G(y)$ to denote the group matrix with y^t as row 0. Note that the i, j element of $\Gamma_G(y)$ is $\gamma_{ij} = \gamma_{0,i^{-1}j} = y_{i^{-1}j}$. We also find it convenient to use the symbol "$\geqslant$" between subscripts. That is, we would like to write $a \geqslant b$ if $a \geqslant b$ in the usual ordering of the real numbers. Caution must be taken, however, since we add subscripts mod n and the order relation is not compatible with addition mod n. Keeping this in mind, we provide the equivalent definition of $U_G(x)$:

$$U_G(x) = (u_{ij}), \qquad \text{where} \quad u_{ij} = \begin{cases} x_{j^{-1}i} & \text{if} \quad j^{-1}i \geqslant i, \\ 0 & \text{otherwise.} \end{cases}$$

The first definition allows us to see how "close" $U_G(x)$ is to a group matrix, whereas the second definition allows us to express the ij element of $U_G(x)\Gamma_G(y)$ precisely. This will be used in the following lemma.

LEMMA 13. *Assume the orthogonality relations relative to G hold for x and y. Then*

$$\mathscr{S}(U_G(x)\Gamma_G(y)) = xy^t.$$

Proof. If $B = U_G(x)\Gamma_G(y)$, then for $i \neq n-1$, the i,j element of $\mathscr{S}(B)$ is

$$b_{ij} - b_{i+1,k_{k+1}^{-1}j} = \sum_{k^{-1}i \geqslant i} x_{k^{-1}i} y_{k^{-1}j} - \sum_{k^{-1}(i+1) \geqslant i+1} x_{k^{-1}(i+1)} y_{k^{-1}k_{i+1}^{-1}j}.$$

Notice that

$$\{k^{-1}i \mid k^{-1}i \geqslant i\} = \{i, i+1, i+2, \ldots, i+(n-i-1)\}$$

and that if $k^{-1}i = i+m$ for some $m \in \{0,1,\ldots,n-i-1\}$, then if $i \neq 0$, $k = \prod_{r=1}^{m} k_{i+r}$. Thus, if we set $g(0) := 0$ and $g(h) := (\prod_{r=1}^{h} k_{i+r})^{-1}$, then we can write $k^{-1} = g(m)$. Similarly, if $k^{-1}(i+1) = i+m$ for some $m \in \{1,2,\ldots,n-i-1\}$, then

$$k^{-1} = k_{i+m}^{-1} k_{i+m-1}^{-1} \cdots k_{i+2}^{-1} = g(m)k_{i+1}.$$

Hence, since $g(m)k_{i+1}(i+1) = g(m)i$ and $g(m)k_{i+1}k_{i+1}^{-1}j = g(m)j$, we have

$$b_{ij} - b_{i+1,k_{i+1}^{-1}j} = \sum_{m=0}^{n-i-1} x_{g(m)i} y_{g(m)j} - \sum_{m=1}^{n-i-1} x_{g(m)k_{i+1}(i+1)} y_{g(m)k_{i+1}k_{i+1}^{-1}j}$$

$$= x_{g(0)i} y_{g(0)j} = x_i y_j,$$

which is the i,j element of xy^t.

If $i = n-1$, then the i,j element of $\mathscr{S}(B)$ is $b_{n-1,j} - b_{0,k_{n-1}^{-1}j}$. We show that $b_{0,k_{n-1}^{-1}j}$ is zero for every $j \in G$. This will imply that the i,j element of $\mathscr{S}(B)$ is

$$b_{n-1,j} = \sum_{k^{-1}(n-1) \geqslant n-1} x_{k^{-1}(n-1)} y_{k^{-1}j} = x_{n-1} y_j$$

and so will finish the proof. For each $j \in G$, recall that E_j denotes the permutation matrix defined by

$$E_j z = (z_{0j}, z_{1^{-1}j}, z_{2^{-1}j}, \ldots, z_{(n-1)^{-1}j})^t \qquad \text{for} \quad z \in R^n.$$

Let $\Gamma_G(y) = (\gamma_{ij})$ and recall that $\gamma_{ij} = y_{i^{-1}j}$, so

$$\Gamma_G(y) = [E_0 y, E_1 y, \ldots, E_{n-1} y].$$

By definition, row 0 of $U_G(x)$ is $(E_0 x)^t = x^t E_0^t$. Let $k \in G$. Then, since $E_0 = E_0^t$,

$$\begin{aligned} b_{0,k} &= x^t E_0^t E_k y \\ &= x^t E_0^t (y_{0k}, y_{1^{-1}k}, \ldots, y_{(n-1)^{-1}k}). \end{aligned}$$

Let $z_m := y_{m^{-1}k}$, so $z_{m^{-1}} = y_{mk}$. Then

$$\begin{aligned} b_{0,k} &= x^t E_0^t z \\ &= x^t (z_0, z_{1^{-1}}, z_{2^{-1}}, \ldots, z_{(n-1)^{-1}})^t \\ &= x^t (y_k, y_{1k}, y_{2k}, \ldots, y_{(n-1)k})^t \\ &= x^t F_k y, \end{aligned}$$

where F_k is the permutation matrix used to develop the orthogonality relations. Hence $b_{0,k} = 0$ for any $k \in G$, which concludes the proof. ■

We may now state the main result of the section.

THEOREM 14. *Let A be an $n \times n$ matrix over R. Let $\mathscr{G}(A) = \sum_{m=1}^{\alpha} x_m y_m^t$. Then A can be expressed as*

$$A = \sum_{m=1}^{\alpha} U_G(x_m)\Gamma_G(y_m) + K,$$

where K is the group matrix with the same first row as A, $\Gamma_G(y_m)$ is a group

matrix with first row y_m^t, *and* $U_G(x_m)$ *is "close" to a group matrix in the sense of Definition* 12.

Proof. By Lemma 13, we know that $\mathscr{G}(A) = \mathscr{G}\left[\sum_{m=1}^{\alpha} U_G(x_m)\Gamma_G(y_m)\right]$. Since $R[G] = \ker \mathscr{G}$, we have $A - \sum_{m=1}^{\alpha} U_G(x_m)\Gamma_G(y_m) = K \in R[G]$. It remains to be shown that the first row of $\sum_{m=1}^{\alpha} U_G(x_m)\Gamma_G(y_m)$ is zero. This can be accomplished using the general orthogonality relations similar to the way the orthogonality conditions were used to show the first row of $U_G(x)\Gamma_G(y)$ was zero in the proof of the previous lemma. ■

As we noted previously, Theorem 14 is equivalent to Corollary 6 for the special case of circulant matrices.

CONCLUSION AND SUGGESTIONS FOR FURTHER RESEARCH

We have shown how arbitrary square matrices over a ring R with identity can be written as sums of products of group matrices and matrices "close" to group matrices. This is accomplished by defining a displacement operator relative to a particular group. The displacement operator can be applied to a square matrix, resulting in a displacement matrix. The displacement matrix can then be expressed as a sum of outer products of vectors. The vectors can be used to construct the desired expression in a straightforward way involving no further computation, as described in the theorems and corollaries in this paper.

Suppose $R = \mathbb{C}$ and a matrix A is expressed as

$$A = \sum_{m=1}^{\alpha} U_m Y_m + K$$

using the SVD of $\mathscr{G}(A)$

$$\mathscr{G}(A) = \sum_{m=1}^{\alpha} \sigma_m u_m v_m^t,$$

where $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_\alpha > 0$. It is well known that if we take the first $k < \alpha$ terms to form the matrix

$$\mathscr{G}(A)_k = \sum_{m=1}^{k} U_m Y_m + K,$$

then $\mathscr{G}(A)_k$ is the best rank k approximation to $\mathscr{G}(A)$ in the least squares sense. Is it also true that

$$A_k = \sum_{m=1}^{k} U_m Y_m + K$$

is the "best" approximation of this form we can obtain for A in some sense?

Some interesting group theoretic questions can also be asked. If G_1 and G_2 are two groups, then one could ask whether there is any relationship between the relative structure of the groups and the structure of the displacement matrices of group matrices for G_1 with respect to the group G_2. For example, consider the case that G_1 is the cyclic group of order $n = 2m$ and D_n is the dihedral group of order n. A typical matrix $B \in R[D_n]$ is of the form

$$B = \left[\begin{array}{c|c} S & T \\ \hline T & S \end{array}\right],$$

where S is circulant and T is constant along the antidiagonals (or a retrocirculant). This reflects the fact that D_n can be viewed as a semidirect product of a cyclic group of order 2 and one of order m [25]. The displacement matrix $\tilde{\mathscr{G}}(B)$ has the form

$$\tilde{\mathscr{G}}(B) = \left[\begin{array}{c|c} W & V \\ \hline V & W \end{array}\right],$$

where

$$W = \begin{bmatrix} & & & & * \\ & & 0 & & * \\ & & & & \vdots \\ & & & & * \\ * & * & \cdots & * & 0 \end{bmatrix},$$

so that $\operatorname{rank}(W) \leqslant 2$. Similarly, a group matrix for a direct sum of cyclic groups is a block circulant with circulant blocks. If A is an $mn \times mn$ block circulant with circulant blocks, then $\tilde{\mathscr{G}}(A)$ is a matrix of blocks having the same structure as W above. Certainly there are some patterns here. Can they be formulated in a consistent way that would provide a method for express-

ing structural differences or similarities in finite groups? Such questions are interesting from an algebraic prospective.

I would like to thank Mary Mikla for typing the manuscript.

REFERENCES

1 P. J. Davis, *Circulant Matrices*, Wiley, New York, 1979.
2 R. Blahut, *Fast Algorithms for Digital Signal Processing*, Addison-Wesley, Reading, Mass., 1985.
3 J. H. McCllelan and C. M. Rader (Eds.), *Number Theory in Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, N.J., 1979.
4 P. D. Gader, Tridiagonal factorizations of Fourier matrices with applications to parallel computation of discrete Fourier transforms, *Linear Algebra Appl.*, 102:169–209 (1988).
5 B. N. Parlett, Winograd's Fourier transform via circulants, *Linear Algebra Appl.* 33:111–122 (1982).
6 D. J. Rose, Matrix identities of the fast Fourier transform, *Linear Algebra Appl.* 29:423–443 (1980).
7 K. Wang, Resultants and group matrices, *Linear Algebra Appl.* 33:111–122 (1980).
8 K. Wang, On the generalization of a retrocirculant, *Linear Algebra Appl.* 37:35–43 (1981).
9 K. Wang, On the group matrices for a generalized dihedral group, *Linear Algebra Appl.* 39:83–89 (1981).
10 R. Chalkley, Information about group matrices, *Linear Algebra Appl.* 38:121–133 (1981).
11 D. Chillag, Generalized circulants and class functions of finite groups, *Linear Algebra Appl.* 93:191–208 (1987).
12 M. Behzad, G. Chartrand, and L. Lesniak Foster, *Graphs and Digraphs*, Wadsworth International Group, Belmont, Calif., 1979.
13 D. Eberly and P. Hartung, Group convolutions and matrix transforms, *SIAM J. Algebraic Discrete Methods* 8:263–275 (1987).
14 P. D. Gader, Image algebra representations of group convolutions on Cayley networks, *SIAM J. Discrete Math.*, submitted for publication.
15 G. Carlsson, J. Cruthirds, H. Sexton, and C. Wright, Interconnection networks based on a generalization of cube-connected cycles, *IEEE Trans. Comput.* C-34:769–772 (1985).
16 F. Preparata and J. Vuillemin, The cube-connected cycles: A versatile network for parallel computation, *Comm. ACM* 24:300–309 (1981).
17 S. Akers and B. Krishnamurphy, Group graphs as interconnection networks, in *Digest of Papers, Fourteenth International Conference on Fault Tolerant Computing*, IEEE Press, 1984.
18 V. Faber, Global communication algorithms for hypercubes and other Cayley coset graphs, *SIAM J. Discrete Math.*, submitted for publication.

19 D. Eberly and D. Wenzel, Adaptation of group algebras to image processing, preprint.
20 T. Kailath, S. Kung, and M. Morf, Displacement ranks of matrices and linear equations, *J. Math. Anal. Appl.* 68:395–407 (1979).
21 T. Kailath, B. Levy, L. Ljung, and M. Morf, The factorization and representation of operators in the algebra generated by Toeplitz operators, *SIAM J. Appl. Math.* 37:467–484 (1979).
22 B. Friedlander, M. Morf, and T. Kailath, New inversion formulas for matrices classified in terms of their distance from Toeplitz matrices, *Linear Algebra Appl.* 27:31–60 (1979).
23 Y. Bistritz and T. Kailath, Inversion and factorization of non-hermitian quasi-Toeplitz matrices, *Linear Algebra Appl.* 98:77–121 (1988).
24 T. Kailath, B. Levy, L. Ljung, and M. Morf, Fast time-invariant implementations of Gaussian signal detectors, *IEEE Trans. Inform. Theory* 24:469–476 (1978).
25 B. Levy, T. Kailath, L. Ljung, and M. Morf, Fast time-invariant implementations for linear least-squares smoothing filters, *IEEE Trans. Automat. Control* 24:770–774 (1979).
26 M. Weinstein, *Examples of Groups*, Polygonal Publishing House, Passaic, N.Y., 1977, pp. 10–17.
27 G. Ammar and P. Gader, A variant of the Gohberg-Semencul formula involving circulant matrices, *SIAM J. Matrix Anal. Appl.*, to appear.
28 G. Ammar and P. Gader, New decompositions of the inverse of a Toeplitz matrix, in *Proceedings of the International Symposium on the Mathematical Theory of Networks and Systems*, Amsterdam, 19–23 June, 1989.

Received 4 April 1988; final manuscript accepted 28 June 1989