THE ALGEBRA GENERATED BY SIMPLE ELEMENTS OF A MATRIX CENTRALIZER

RALPH JOHN DE LA CRUZ

(Communicated by H. Radjavi)

Abstract. Let $\mathscr{C}(S)$ denote the centralizer of an arbitrary square matrix *S*. An element $A \in \mathscr{C}(S)$ is simple if A - I is of rank 1. Let \mathscr{A}_S denote the subalgebra generated by the simple elements of $\mathscr{C}(S)$. We use the Weyr canonical form to describe the subalgebra \mathscr{A}_S , and we show that if $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of *S*, and *l* is the number of defective eigenvalues of *S*, then \mathscr{A}_S is of dimension $l + \sum_{i=1}^k \text{nullity}(S - \lambda_i I)^2$.

1. Introduction

We consider matrices over an algebraically closed field \mathbb{F} with zero characteristic. Let \mathscr{S} be any set of *n*-by-*n* matrices. We call an element $A \in \mathscr{S}$ simple if A - I is of rank 1. The following are known matrix decompositions with simple elements as factors.

- Every *n*-by-*n* matrix with determinant ± 1 is a product of 2n 1 involutions which are simple [10].
- Every *n*-by-*n* orthogonal matrix is a product of n+1 simple orthogonal matrices [11, 14].
- Every 2n-by-2n symplectic matrix is a product of 2n + 1 simple symplectic matrices [2, 5].

For nonsingular matrices A and S, we say that A is ϕ_S -orthogonal if $SA^TS^{-1} = A^{-1}$, or equivalently,

$$AS = SA^{-T}$$

Notice that if S = I, then a ϕ_S -orthogonal matrix is an orthogonal matrix, and that if $S = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$, then a ϕ_S -orthogonal matrix is a symplectic matrix. Thus the ϕ_S -orthogonal matrices may be viewed as generalizations of symplectic and orthogonal matrices. Let \mathcal{O}_S be the set of ϕ_S -orthogonal matrices. If $\mathbb{F} = \mathbb{C}$, then every element in \mathcal{O}_S is a product of simple elements in \mathcal{O}_S if and only if $S^{-T}S$ is an involution [4]. A

Keywords and phrases: Simple, Weyr structure, Weyr canonical form, centralizer, defective.



Mathematics subject classification (2020): 15A03, 15A27.

related study [1] has been recently done for the set of ψ_S -orthogonal matrices, where given nonsingular matrices *S* and *A*, we say that *A* is ψ_S -orthogonal if $S\overline{A^{-1}}S^{-1} = A^{-1}$, or equivalently,

$$AS = S\overline{A}$$

It is shown in [1] that if $\mathbb{F} = \mathbb{C}$, then every ψ_S -orthogonal matrix is a product of simple ψ_S -orthogonal matrices if and only if *S* is consimilar to a diagonal matrix.

For an arbitrary square matrix S the *centralizer* $\mathscr{C}(S)$ of S is the set of all A such that

AS = SA.

If S is a nontrivial involution, and oftentimes assumed to be also Hermitian, the elements of $\mathscr{C}(S)$ are also called S-symmetric, and has been characterized in [12]. Generalizations and eigenvalue problems relating to S-symmetric matrices have also been considered [3, 7, 8, 13]. In this paper, we consider an arbitrary square matrix S and we use the Weyr canonical form to describe the subalgebra generated by the simple elements of $\mathscr{C}(S)$. We use the preceding to prove our main result, which we state in the following theorem. Recall that a *defective eigenvalue* is an eigenvalue whose geometric multiplicity is strictly less than its algebraic multiplicity.

THEOREM 1. Let S be an arbitrary square matrix over an algebraically closed field of zero characteristic, and \mathscr{A}_S be the subalgebra generated by the elements X that satisfy XS = SX and rank(X - I) = 1. Suppose that $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of S and l is the number of defective eigenvalues of S. Then

$$\dim \mathscr{A}_S = l + \sum_{i=1}^k nullity(S - \lambda_i I)^2.$$

The following is immediate from Theorem 1 as we will see in our discussions.

COROLLARY 1. Following the assumptions and notations in Theorem 1, we have that $\mathscr{C}(S) = \mathscr{A}_S$ if and only if the Jordan structure of S corresponding to each eigenvalue is of the form (2, 1, ..., 1) or (1, 1, ..., 1).

In other words, the simple elements of $\mathscr{C}(S)$ generate the whole algebra if and only if *S* is *almost diagonalizable*.

2. Proof of the main result

If $XS_1X^{-1} = S_2$ for some nonsingular matrix X, then

$$\mathscr{C}(S_1) = X^{-1}\mathscr{C}(S_2)X = \{X^{-1}AX \mid A \in \mathscr{C}(S_2)\},\$$

that is, there is an isomorphism between the algebras $\mathscr{C}(S_1)$ and $\mathscr{C}(S_2)$. Thus, to count the dimension of \mathscr{A}_S , we can assume without loss of generality that *S* is in a canonical form under similarity. Both the Jordan Canonical form and Weyr Canonical form imply that if $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of *S*, then we can assume $S = W_1 \oplus \cdots \oplus W_k$ where λ_i is the only eigenvalue of W_i . Due to Sylvester's Theorem [6, Theorem 4.4.6], if *X* commutes with *S*, then $X = X_1 \oplus \cdots \oplus X_k$, where X_i and W_i have the same sizes for i = 1, ..., k. Thus, dim $\mathscr{A}_S = \sum_{i=1}^k \dim \mathscr{A}_{W_i}$. Since $\mathscr{A}_{W_i} \leq \mathscr{C}(W_i)$, we have that $\mathscr{A}_S = \mathscr{C}(S)$ if and only if dim $\mathscr{A}_{W_i} = \dim \mathscr{C}(W_i)$ for i = 1, ..., k.

PROPOSITION 1. Theorem 1 and Corollary 1 are true if they are true for the case when *S* has only one eigenvalue.

If $S = \lambda I_n$ for some $\lambda \in \mathbb{F}$, then $\mathscr{C}(S) = \mathbb{F}^{n \times n}$. Define $E_{i,j}$ to be the matrix whose (i, j)-entry is 1 and whose other entries are 0. Observe that if $i \neq j$, then $E_{i,j} = (I + (2E_{i,j})) - (I + E_{i,j})$ is a difference of simple elements in $\mathscr{C}(S)$. If i = j, then $E_{i,i} = \text{diag}(I_{i-1}, 3, I_{n-i}) - \text{diag}(I_{i-1}, 2, I_{n-i})$ is a difference of simple elements in $\mathscr{C}(S)$. Since the $E_{i,j}$'s form a basis for $\mathbb{F}^{n \times n}$, we have $\mathscr{A}_S = \mathbb{F}^{n \times n}$ and so

 $\dim \mathscr{A}_S = \dim \mathbb{F}^{n \times n} = \operatorname{nullity}(S - \lambda I)^2.$

We are left to prove Theorem 1 for defective eigenvalues.

Let *J* be an *n*-by-*n* Jordan matrix with only one eigenvalue and suppose (m_1, \ldots, m_s) is the Jordan structure of *J*, where $m_1 > m_2 > \cdots > m_s$. Then an *n*-by-*n* blocked matrix $K = [K_{i,j}]$, where each $K_{i,j}$ is m_i -by- m_j , commutes with *J* if and only if $K_{i,j} = \begin{bmatrix} 0 \ T \end{bmatrix}$ for $i \ge j$, and $K_{i,j} = \begin{bmatrix} T \\ 0 \end{bmatrix}$ for $i \le j$, where *T* is a matrix of the form

a_1	a_2	a_3		 a_{m_j}	
0	a_1	a_2	a_3		
0	0	a_1	a_2		
:			۰. ِ		
0	0	0		a_1	
LO	0	0	• • •	u_1	

see [9, Proposition 3.1.2]. For instance, if

$$J = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

then K commutes with J if and only if

$$K = \begin{bmatrix} a \ b \ c \ d \ e \ f \ g \\ 0 \ a \ b \ 0 \ d \ e \ 0 \\ 0 \ 0 \ a \ 0 \ 0 \ d \ 0 \\ \hline h \ i \ j \ k \ l \ m \ n \\ 0 \ h \ i \ 0 \ k \ l \ 0 \\ \hline 0 \ 0 \ h \ 0 \ 0 \ k \ 0 \\ \hline \hline 0 \ 0 \ 0 \ 0 \ p \ q \end{bmatrix}$$

The Frobenius formula [9, Proposition 3.1.3] gives the dimension of $\mathscr{C}(J)$ as $m_1 + 3m_2 + \cdots + (2s-1)m_s$, which justifies the use of 17 variables in the above matrix *K*. We note that the rank of K - I is not immediately obtained from this form, and so we turn to the Weyr canonical form.

A very good reference material for properties and applications of the Weyr canonical form is [9]. We recall some concepts and adapt notations from this book.

DEFINITION 1. (Definition 2.1.1 in [9]) A basic Weyr matrix with eigenvalue λ is an *n*-by-*n* matrix *W* of the following form: There is a partition $n_1 + \cdots + n_r = n$ of *n* with $n_1 \ge \cdots \ge n_r \ge 1$ such that, when *W* is viewed as an *r*-by-*r* blocked matrix $(W_{i,j})$, where the (i, j) block $W_{i,j}$ is an n_i -by- n_j matrix, the following three features are present:

- 1. The main diagonal blocks $W_{i,i}$ are the n_i -by- n_i scalar matrices λI for i = 1, ..., r.
- 2. The first superdiagonal blocks $W_{i,i+1}$ are full column-rank n_i -by- n_{i+1} matrices in reduced row echelon form (that is, an identity matrix followed by zero rows) for i = 1, ..., r 1.
- 3. All other blocks of *W* are zero.

In this case, we say that W has Weyr structure (n_1, \ldots, n_r) . A matrix W is a Weyr matrix, or is in Weyr form if it is a direct sum of basic Weyr matrices with distinct eigenvalues.

We also have that the number n_1 is the nullity of $W - \lambda I_n$. For example,

is a basic Weyr matrix with Weyr structure (4,3,3,1) such that nullity $(W - \lambda I) = 4$.

THEOREM 2. (Theorem 2.2.4 in [9]) To within permutation of basic Weyr blocks, each square matrix A over an algebraically closed field is similar to a unique Weyr matrix W. The matrix W is called the Weyr canonical form of A.

With Proposition 1 in mind, we prove Theorem 1 only for the case when S = W is a nonscalar basic Weyr matrix. The following completely describes the elements in $\mathscr{C}(W)$.

LEMMA 1. (Proposition 2.3.3 in [9]) Let W be an n-by-n basic Weyr matrix with Weyr structure $(n_1, ..., n_r)$, $r \ge 2$. Let K be an n-by-n matrix, blocked according to the partition $n = n_1 + \cdots + n_r$, and let $K_{i,j}$ denote its (i, j) block (an n_i -by- n_j matrix) for i, j = 1, ..., r. Then W and K commute if and only if K is a block upper triangular matrix for which

$$K_{i,j} = \begin{bmatrix} K_{i+1,j+1} * \\ 0 * \end{bmatrix} \text{ for } 1 \leq i \leq j \leq r-1.$$

$$\tag{2}$$

Here, we have written $K_{i,j}$ *as a blocked matrix where the zero block is* $(n_i - n_{i+1})$ *-by-* n_{j+1} *.*

In particular, a matrix in C(W) is completely determined by its top row of blocks. For illustration, suppose a basic Weyr matrix W has Weyr structure (3,3,2,1). When we write $K \in \mathscr{C}(W)$ as

$$K = \begin{bmatrix} 1 & 0 & 1 & 1 & -1 & 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 0 & 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 2 & 0 & 0 & -2 & 1 & 1 & 0 \end{bmatrix}$$

we basically mean that

$$K = \begin{bmatrix} 1 & 0 & 1 & 1 & -1 & 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 0 & 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 2 & 0 & 0 & -2 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

For convenience, a linear combination of elements in $\mathscr{C}(W)$ will also be written simply as a linear combination of the first row of these elements.

THEOREM 3. Let W be a basic Weyr matrix with Weyr structure $(n_1, ..., n_r)$. Let $K = [K_{i,j}] \in \mathscr{C}(W)$ such that $K_{i,j}$ is n_i -by- n_j . Then K is simple if and only if the following hold:

(a)
$$K_{1,1} = \begin{bmatrix} I_{n_2} & B_1 \\ 0 & C_1 \end{bmatrix}$$
.
(b) $K_{1,i} = \begin{bmatrix} 0_{n_2 \times n_{i+1}} & B_i \\ 0 & C_i \end{bmatrix}$ for $i = 2, ..., r - 1$
(c) $K_{1,r} = \begin{bmatrix} B_r \\ C_r \end{bmatrix}$, where B_r has n_2 rows.
(d) $\begin{bmatrix} B_1 & B_2 & ... & B_r \\ C_1 - I & C_2 & ... & C_r \end{bmatrix}$ has rank 1.

Proof. Let $K \in \mathscr{C}(W)$. It is straightforward to verify that if K satisfies the set of conditions above, then K is simple. Now, suppose K is simple. In view of Lemma 1, if $K_{i,i} \neq I$ for some i = 2, ..., r, then $K_{1,1} = \begin{bmatrix} K_{i,i} & * \\ 0 & * \end{bmatrix} \neq I$ and we have $\begin{pmatrix} \begin{bmatrix} K_{1,1} - I & * \\ & \ddots & \end{bmatrix} \end{pmatrix}$

$$\operatorname{rank}(K-I) = \operatorname{rank}\left(\left| \begin{array}{c} \ddots \\ K_{i,i} - I \\ 0 \\ K_{r,r} - I \end{array} \right| \right)$$
$$\geq \operatorname{rank}(K_{i,i} - I) + \operatorname{rank}(K_{1,1} - I)$$
$$\geq 2.$$

This is a contradiction since K - I is of rank 1, and so $K_{i,i} = I$ for i = 2, ..., r, that is,

$$K = \begin{bmatrix} K_{1,1} \dots & & \\ & I_{n_2} & K_{2,3} & \\ & \ddots & \ddots & \\ & & & I_{n_{r-1}} & K_{r-1,r} \\ 0 & & & & I_{n_r} \end{bmatrix}$$

Suppose that $K_{i,i+j} \neq 0$ for some $2 \leq i \leq r-1$ and j = 1, ..., r-i. We assume without loss of generality that $K_{i,i+j}$ is the immediate nonzero superdiagonal block. This gives us,

$$\operatorname{rank}(K-I) = \operatorname{rank}\left(\begin{bmatrix} K_{1,1} - I \cdots K_{1,1+j} & & \\ & 0 & 0 & \ddots & \\ & & \ddots & 0 & K_{i,i+j} \\ & & 0 & \ddots & \\ & & & \ddots & 0 \\ 0 & & & & 0 \end{bmatrix} \right)$$

$$\geq \operatorname{rank}(K_{1,1+j}) + \operatorname{rank}(K_{i,i+j})$$

$$\geq 2,$$

which is again a contradiction to the assumption that *K* is simple. This implies that $K = \begin{bmatrix} K_{1,1} & * \\ 0 & I \end{bmatrix}$ and so the rank of K - I is equal to the rank of the first row of blocks of K - I. The previous discussion and Lemma 1 give us the desired form of *K*. \Box

The characterization in Theorem 3 of simple elements in $\mathscr{C}(W)$ implies that \mathscr{A}_W is contained in the algebra \mathscr{B}_W of matrices in $\mathscr{C}(W)$ whose first row of blocks is of the form

$$K = \begin{bmatrix} \alpha I_{n_2} & B_1 \\ 0 & C_1 \end{bmatrix} \begin{bmatrix} 0_{n_2 \times n_{i+1}} & B_2 \\ 0 & C_2 \end{bmatrix} \cdots \begin{bmatrix} B_r \\ C_r \end{bmatrix},$$
(3)

where each $B_i \in \mathbb{F}^{n_2 \times (n_i - n_{i+1})}$ and each $C_i \in \mathbb{F}^{(n_1 - n_2) \times (n_i - n_{i+1})}$ for $i = 1, \ldots, r-1$, $B_r \in \mathbb{F}^{n_2 \times n_r}$, and $C_r \in \mathbb{F}^{(n_1 - n_2) \times n_r}$. It is then straightforward to count that the dimension of \mathscr{B}_W is

$$1 + n_1((n_1 - n_2) + (n_2 - n_3) + \dots + (n_{r-1} - n_r) + n_r) = 1 + n_1^2.$$

Hence, since $n_1 = \text{nullity}(W - I)$, we are done if we show that $\mathscr{B}_W = \mathscr{A}_W$. It is enough to prove that the matrix *K* in Equation 3 is in \mathscr{A}_W , for the cases when

- $K = \alpha I$ for some $\alpha \in \mathbb{F}$, or
- $\alpha = 0$ and all other blocks are zero except for one which is a matrix of the form $E_{k,l}$.

Indeed, $I \in \mathscr{A}_W$ since it is the linear combination

$$2\left[\begin{bmatrix}I_{n_2} & E_{i,j} \\ 0 & I_s\end{bmatrix} 0\right] - \left[\begin{bmatrix}I_{n_2} & 2E_{i,j} \\ 0 & I_s\end{bmatrix} 0\right]$$
(4)

of simple elements in $\mathscr{C}(W)$, and so $K = \alpha I \in \mathscr{A}_W$ for all α . Also if K satisfies the other condition, one checks that K + I is a simple element, and thus $K = (K + I) - I \in \mathscr{A}_W$. This proves Theorem 1.

REMARK 1. In both the defective and nondefective case, we have shown that a basis containing only simple elements exists for \mathscr{A}_S .

We now consider Corollary 1 for the case when *S* has only one eigenvalue. If $S = \lambda I$, that is, when the Jordan structure of *S* is $(1, \ldots, 1)$, we have shown that $\mathscr{A}_S = \mathscr{C}(S)$. Assume that *S* is a nonscalar basic Weyr matrix with eigenvalue λ . We use Proposition 3.2.2 in [9] which implies that if *S* has Weyr structure (n_1, \ldots, n_r) , the dimension of $\mathscr{C}(S)$ is $n_1^2 + \cdots + n_r^2$. Thus, Theorem 1 and the fact that $n_1 = \text{nullity}(S - \lambda I)$ imply that $\mathscr{A}_S = \mathscr{C}(S)$ if and only if $1 = n_2^2 + \cdots + n_r^2$, and this happens only when r = 2 and $n_2 = 1$. In summary, if *S* is a nonscalar basic Weyr matrix with eigenvalue λ , $\mathscr{A}_S = \mathscr{C}(S)$ if and only if the Weyr structure of *S* is $(n_1, 1)$. One easily checks that a basic Weyr matrix with this Weyr structure has Jordan structure $(2, 1, \ldots, 1)$ (or one may consult Corollary 2.4.3 in [9] for the relation of the Weyr structure and Jordan structure).

REFERENCES

- [1] T. AGAPITO, A. T. PARAS, ψ_S -orthogonal matrices and ψ_S -symmetries, Linear Algebra Appl. **584** (2020) 185–196.
- [2] R. J. DE LA CRUZ, K. DELA ROSA, D. I. MERINO, A. T. PARAS, *The Cartan-Diedonne-Scherk Theorems for complex S-orthogonal matrices*, Linear Algebra Appl. 458 (2014) 251–260.
- [3] M. CATRAL, L. LEBTAHI, J. STUART, N. THOME, On a matrix group constructed from an $\{R, s+1, k\}$ -potent matrix, Linear Algebra Appl. **461** (2014) 200–210.
- [4] R. J. DE LA CRUZ, D. I. MERINO, A. T. PARAS, S orthogonal matrices and S symmetries, Linear Algebra Appl. 474 (2015) 213–229.
- [5] J. DIEUDONNE, Sur les generateurs des groupes classiques, Summa Bras. Math. 3 (1955) 149–178.

- [6] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.
- [7] G. HUANG, F. YIN, Constrained inverse eigenproblem and associated approximation problem for anti-Hermitian R-symmetric matrices, Applied Mathematics and Computation 186 (2007) 426–434.
- [8] G. HUANG, F. YIN, Matrix inverse problem and its optimal approximation problem for R-symmetric matrices, Applied Mathematics and Computation 189 (2007) 482–489.
- [9] K. O'MEARA, J. CLARK, C. VINSONHALER, Advanced topics in linear algebra: weaving matrix problems through the Weyr form, Oxford University Press, Oxford 2011.
- [10] H. RADJAVI, Decomposition of matrices into simple involutions, Linear Algebra Appl., 12 (1975) 247–255.
- P. SCHERK, On the decomposition of orthogonalities into symmetries, Proc. Amer. Math. Soc. 1 (1950) 481–491.
- [12] W. TRENCH, Characterization and properties of matrices with generalized symmetry or skew symmetry, Linear Algebra Appl. 377 (2004) 207–218.
- [13] W. TRENCH, Minimization problems for (R,S)-symmetric and (R,S)-skew symmetric matrices, Linear Algebra Appl. **389** (2004) 23–31.
- [14] F. UHLIG, Constructive ways for generating (generalized) real orthogonal matrices as products of (generalized) symmetries, Linear Algebra Appl. 332–334 (2001) 459–467.

(Received March 27, 2020)

Ralph John de la Cruz Institute of Mathematics University of the Philippines Diliman, Quezon City, Philippines e-mail: rjdelacruz@math.upd.edu.ph

Operators and Matrices www.ele-math.com oam@ele-math.com