Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

THE PRODUCTS OF INVOLUTIONS IN A MATRIX CENTRALIZER*

RALPH JOHN DE LA CRUZ† AND RAYMOND LOUIS TAÑEDO†

Abstract. A square matrix A is an *involution* if $A^2 = I$. The *centralizer* of a square matrix S denoted by $\mathscr{C}(S)$ is the set of all A such that $AS = SA$ over an algebraically closed field of characteristic not equal to 2. We determine necessary and sufficient conditions for $A \in \mathscr{C}(S)$ to be a product of involutions in $\mathscr{C}(S)$ where S is a basic Weyr matrix with homogeneous Weyr structure of length 3. Finally, we will show some results for the case when the length of the Weyr structure is greater than 3.

Key words. Involution, Weyr form, Centralizer, Matrix decompositions.

AMS subject classifications. 15A21, 15A23, 15A27.

1. Introduction. A nonsingular matrix A is said to be an *involution* if $A^2 = I$. It is known that a complex matrix A is a product of involutions if and only if $\det A = \pm 1$. Gustafson et al. [1] showed that the number of factors may be taken to be four, that is, every square complex matrix A with determinant ± 1 may be written as

$$A = \prod_{i=1}^{4} A_i = A_1 A_2 A_3 A_4, \tag{1.1}$$

where A_1, A_2, A_3, A_4 are involutions. Fix a square matrix S that has the same size as A. If each A_i in (1.1) commutes with S, then it is easy to see that A commutes with S. But is the converse true? Let $\mathbb{F}$ be an algebraically closed field such that char $\mathbb{F} \neq 2$. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and set $S = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ for some $\lambda \in \mathbb{F}$. A matrix X commutes with S if and only if $X = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$ for some $a, b \in \mathbb{F}$. In particular, A commutes with S. Now the matrix X is an involution if and only if $X = \pm I_2$. Hence A, which has determinant 1, is not a product of involutions that commutes with S.

Denote by $\mathscr{C}(S)$ the set of all matrices that commute with S, which is also called the *centralizer* of S. We are now interested in the following problem: For a fixed square matrix S and a matrix $A \in \mathscr{C}(S)$, what condition/s on A is/are needed for A, aside from $\det A = \pm 1$, for A to be written as product as in (1.1), where each $A_i \in \mathscr{C}(S)$? In Section 2, we show that our problem is invariant under similarity of S. We discuss briefly the Weyr Canonical Form, which is the preferred form over the Jordan form when it comes to problems concerning matrix centralizers. We then show that we can further reduce the problem to the case when S is a basic Weyr matrix and so we turn our focus to this case. In this paper, we consider S a basic Weyr matrix with homogeneous Weyr structure of length r. The following is the case when S has length at most three.

*Received by the editors on April 21, 2022. Accepted for publication on July 23, 2022. Handling Editor: João Filipe Queiró. Corresponding Author: Ralph John de la Cruz

†Institute of Mathematics, University of the Philippines, Natural Sciences Research Institute, Quezon City 1101, Philippines (rjdelacruz@math.upd.edu.ph, rltanedo@math.upd.edu.ph).

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

THEOREM 1.1. *Let S be a basic Weyr matrix with homogeneous Weyr structure of length r.*

(a) *Suppose $r = 2$. A matrix X is a product of involutions in $\mathscr{C}(S)$ if and only if*

$$X = \begin{bmatrix} X_1 & X_2 \\ 0 & X_1 \end{bmatrix},$$

with $\det X_1 = \pm 1$ *and* $\mathrm{tr}(X_1^{-1} X_2) = 0$.

(b) *Suppose $r = 3$. A matrix X is a product of involutions in $\mathscr{C}(S)$ if and only if*

$$X = \begin{bmatrix} X_1 & X_2 & X_3 \\ 0 & X_1 & X_2 \\ 0 & 0 & X_1 \end{bmatrix},$$

with $\det X_1 = \pm 1$, $\mathrm{tr}(X_1^{-1} X_2) = 0$, *and* $2\mathrm{tr}(X_1^{-1} X_3) = \mathrm{tr}((X_1^{-1} X_2)^2)$.

We also study the general case and we illustrate why a finite list of necessary and sufficient conditions may not exist for a matrix to be a product of involutions in $\mathscr{C}(S)$ where S is a basic Weyr matrix with homogeneous Weyr structure of length r for any r.

2. Preliminaries. In this paper, we consider an algebraically closed field such that char $\mathbb{F} \neq 2$. Denote $M_n(\mathbb{F})$ the set of $n \times n$ matrices over $\mathbb{F}$. We first prove some preliminary result.

THEOREM 2.1. *Suppose $S_1, S_2 \in M_n(\mathbb{F})$ such that $S_1 = X^{-1} S_2 X$ for some nonsingular matrix X.*

(a) *A matrix $B \in M_n(\mathbb{F})$ belongs to $\mathscr{C}(S_2)$ if and only if $A = X^{-1} B X$ belongs to $\mathscr{C}(S_1)$.*
(b) *Then A is a product of involutions in $\mathscr{C}(S_1)$ if and only if B is a product of involutions in $\mathscr{C}(S_2)$.*

Proof.

(a) Assume that $A \in \mathscr{C}(S_1)$ such that $A = X^{-1} B X$. Since $AS_1 = S_1 A$, then $(X^{-1} B X)(X^{-1} S_2 X) = (X^{-1} S_2 X)(X^{-1} B X)$ and hence,

$$X^{-1} B S_2 X = X^{-1} S_2 B X.$$

Thus, $S_2 B = B S_2$ and therefore, $B \in \mathscr{C}(S_2)$. The converse is proven similarly.

(b) Assume that $A = \prod_i A_i$ such that $A_i^2 = I$ and $A_i \in \mathscr{C}(S_1)$. By Theorem 2.1 a), $B_i := X^{-1} A_i X \in \mathscr{C}(S_2)$ and $(B_i)^2 = (X^{-1} A_i X)^2 = I$, that is, B_i is an involution for all i. Then $B = X^{-1} A X$ implies that

$$\begin{aligned} B &= X^{-1} \left(\prod_i A_i \right) X = X^{-1} A_1 A_2 \cdots X \\ &= (X^{-1} A_1 X)(X^{-1} A_2 X) \cdots = B_1 B_2 \cdots \\ &= \prod_i B_i. \end{aligned}$$

The converse is proven similarly. ∎

Theorem 2.1(b) tells us that a canonical form of matrices under similarity can be considered instead of an arbitrary matrix S to study products of involutions in $\mathscr{C}(S)$. One of the known canonical forms of

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

matrices is the Jordan canonical form. Let J be the Jordan form of $S \in M_n(\mathbb{F})$ with Jordan structure $(m_1, m_2, \ldots, m_r)$ and let K be an n-by-n matrix, blocked conformally to J, and $K_{i,j}$ denote its (i,j) block for all $i, j = 1, 2, \ldots, r$. By [3, Proposition 3.1.2], J commutes with K if and only if for each of the r^2 blocks $K_{i,j}$, the following hold:

- For $i \geq j$, each $K_{i,j}$ is of the form

$$K_{i,j} = \begin{bmatrix} 0 & \cdots & 0 & a & \cdots & x & y & z \\ 0 & & & & a & \cdots & x & y \\ 0 & & & & & \ddots & & x \\ \vdots & & & & & & & \vdots \\ 0 & \cdots & 0 & & & \cdots & & a \end{bmatrix},$$

  with the first $m_j - m_i$ columns are zero; and
- for $i \leq j$,

$$K_{i,j} = \begin{bmatrix} a & b & c & \cdots & \cdots & z \\ 0 & a & b & c & \cdots & \\ 0 & 0 & a & b & & \\ \vdots & & & \ddots & & \\ 0 & 0 & 0 & \cdots & & a \\ 0 & 0 & 0 & \cdots & & 0 \\ \vdots & \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \cdots & & 0 \end{bmatrix}.$$

Thus, if we consider the Jordan form J with Jordan structure structure $(3,3,2)$, we have that $K \in \mathscr{C}(J)$ if and only if K is a matrix of the form

$$K = \left[\begin{array}{ccc|ccc|cc} a & b & c & j & k & \ell & m & n \\ 0 & a & b & 0 & j & k & 0 & m \\ 0 & 0 & a & 0 & 0 & j & 0 & 0 \\ \hline p & q & r & d & e & f & w & x \\ 0 & p & q & 0 & d & e & 0 & w \\ 0 & 0 & p & 0 & 0 & d & 0 & 0 \\ \hline 0 & s & t & 0 & u & v & g & h \\ 0 & 0 & s & 0 & 0 & u & 0 & g \end{array}\right].$$

Notice that it is hard to determine involutions in $\mathscr{C}(J)$ from this form. Hence, it is not efficient to use the Jordan Canonical form of a matrix to study products of involutions in $\mathscr{C}(S)$ for an arbitrary matrix S. Instead, we use a canonical form of matrices under similarity for identifying products of involutions called the Weyr canonical form whose centralizer is 'nicer' than that of the Jordan Canonical Form.

Definition 2.2 ([3]). *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 $n_1 \geq \cdots \geq n_r \geq 1$ 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:*

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

1. *The main diagonal blocks $W_{i,i}$ are the n_i-by-n_i scalar matrices λI for $i = 1, \ldots, r$.*
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 $i = 1, \ldots, r-1$.*
3. *All other blocks of W are zero.*

In this case, we say that W has Weyr structure $(n_1, n_2, \ldots, n_r)$ *of length 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. In particular, a Weyr matrix with one eigenvalue is a basic Weyr matrix.*

Let W be a basic Weyr matrix with Weyr structure $(n_1, n_2, \ldots, n_r)$. Then, W is said to have a *homogeneous* Weyr structure if $n_1 = n_2 = \cdots = n_r$. Otherwise, W has a *nonhomogeneous* Weyr structure.

Example 2.3. Consider a basic Weyr matrix W with eigenvalue λ of Weyr structure $(5, 3, 2, 1)$. Then W is of the form

$$W = \left[\begin{array}{ccccc|ccc|cc|c}
\lambda & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & \lambda & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & \lambda & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & \lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \lambda & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & \lambda & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \lambda & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda & 0 & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda
\end{array}\right].$$

Theorem 2.4 ([3, Theorem 2.2.4]). *Up to 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.*

Let S be an n×n matrix over $\mathbb{F}$. By Theorem 2.4, S is similar to a unique Weyr matrix W, that is $S = X^{-1}WX$ for some nonsingular matrix X. Let $A \in \mathscr{C}(S)$. By Theorem 2.1(a), $B = X^{-1}AX \in \mathscr{C}(W)$ and so by Theorem 2.1(b), A is a product of involutions in $\mathscr{C}(S)$ if and only if B is a product of involutions in $\mathscr{C}(W)$. From this, it is enough to consider the products of involutions in $\mathscr{C}(W)$.

We use the next theorem to further reduce our problem.

Theorem 2.5 ([2, Sylvester's Theorem]). *Let $\mathbb{F}$ be an algebraically closed field. Let A and B be n-by-n and m-by-m matrices over $\mathbb{F}$, respectively. If A and B have no eigenvalues in common, then for each $n \times m$ matrix C, the equation $AX - XB = C$ has a unique solution $X \in M_{n,m}(\mathbb{F})$. In particular, $AX - XB = 0$ only has the trivial solution.*

Recall that if $A = \bigoplus A_i$, then A is an involution if and only if A_i is an involution for all i. We also describe a matrix in $\mathscr{C}(W)$ as a product of involutions in $\mathscr{C}(W)$.

Theorem 2.6. *Let $\lambda_1, \lambda_2, \ldots, \lambda_r \in \mathbb{F}$ be distinct and let W_i be a basic Weyr matrix corresponding to the eigenvalue λ_i. Suppose that $W = \bigoplus_{i=1}^{r} W_i$.*

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

(a) *A matrix $K \in M_n(\mathbb{F})$ belongs to $\mathscr{C}(W)$ if and only if $K = \bigoplus_{i=1}^{r} K_i$ where $K_i \in \mathscr{C}(W_i)$ for each $i = 1, 2, \ldots, r$.*

(b) *Let $Z = \operatorname{diag}(Z_1, Z_2, \ldots, Z_r) \in \mathscr{C}(W)$. Then Z is a product of involutions in $\mathscr{C}(W)$ if and only if Z_i is a product of involutions in $\mathscr{C}(W_i)$ for each $i = 1, 2, \ldots, r$.*

Proof.

(a) Assume that K belongs to $\mathscr{C}(W)$. Partition K conformally to W. Then, $K = [K_{i,j}]$ is a block matrix for all $i, j = 1, 2, \ldots, r$. Since $KW = WK$, then for $i \neq j$, $W_iK_{i,j} = K_{i,j}W_j$ which implies that $W_iK_{i,j} - K_{i,j}W_j = 0$. By Theorem 2.5, $K_{i,j} = 0$ for $i \neq j$, since W_i and W_j have no eigenvalues in common. Hence, $K = \bigoplus_{i=1}^{r} K_{ii}$. Now, $KW = WK$ implies that for $i = j$ and $\lambda_i = \lambda_j$, $W_iK_{ii} = K_{ii}W_i$ since $KW = WK$. Therefore, $K_{ii} \in \mathscr{C}(W_i)$, for all $i = 1, 2, \ldots, r$. The converse is easily proven.

(b) Assume that $Z = \prod_j \tilde{Z}_j$, where $\tilde{Z}_j$ is an involution in $\mathscr{C}(W)$. By Theorem 2.6 a), $\tilde{Z}_j = \bigoplus_i \tilde{Z}_{j,i}$ where $\tilde{Z}_{j,i} \in \mathscr{C}(W_i)$. Since $\tilde{Z}_j$ is an involution, then $\tilde{Z}_{j,i}$ is an involution in $\mathscr{C}(W_i)$. Thus, $Z_i = \prod_j \tilde{Z}_{j,i}$ and therefore, Z_i is a product of involutions in $\mathscr{C}(W_i)$.
Conversely, assume that $Z_i = \prod_{k=1}^{n_i} Z_{k,i}$ where $Z_{k,i}$ is an involution in $\mathscr{C}(W_i)$ for all $i = 1, 2, \ldots, r$. Let $\beta = \max\{n_i \mid 1 \leq i \leq r\}$. If $n_i < \beta$, then define $Z_{\ell,i} = I$ for all $\ell > n_i$. Hence, for $\ell = 1, 2, \ldots, \beta$, $\tilde{Z}_\ell := \bigoplus_{i=1}^{r} Z_{\ell,i}$ is an involution in $\mathscr{C}(W)$. Thus, $Z = \prod_{\ell=1}^{\beta} \tilde{Z}_\ell$ is a product of involutions in $\mathscr{C}(W)$. $\square$

Theorem 2.6(b) tells us that without loss of generality, we can further assume that S is a basic Weyr matrix. The following describes the elements in the centralizer of a basic Weyr matrix.

LEMMA 2.7 ([3, Proposition 2.3.3]). *Let W be an n-by-n basic Weyr matrix with Weyr structure $(k_1.k_2, \ldots, k_r)$, $r \geq 2$. Let X be an n-by-n matrix, blocked according to the partition $n = k_1 + k_2 + \cdots + k_r$, and let $X_{i,j}$ denote its (i,j) block (an n_i-by-n_j matrix) for $i, j = 1, \ldots, r$. Then W and X commute if and only if X is a block upper triangular for which*

$$X_{i,j} = \begin{bmatrix} X_{i+1,j+1} & * \\ 0 & * \end{bmatrix} \text{ for } 1 \leq i \leq j \leq r-1, \tag{2.2}$$

where the column of asterisks disappears if $n_j = n_{j+1}$ and the $\begin{bmatrix} 0 & * \end{bmatrix}$ row disappears if $n_i = n_{i+1}$.

Example 2.8. Suppose W is a basic Weyr matrix with Weyr structure $(3, 2, 2, 1)$. By Lemma 2.7, a matrix X that commutes with W is of the form

$$X = \left[\begin{array}{ccc|cc|cc|c} a & b & d & g & i & k & m & p \\ 0 & c & e & h & j & \ell & n & q \\ 0 & 0 & f & 0 & 0 & 0 & o & r \\ \hline 0 & 0 & 0 & a & b & g & i & k \\ 0 & 0 & 0 & 0 & c & h & j & \ell \\ \hline 0 & 0 & 0 & 0 & 0 & a & b & g \\ 0 & 0 & 0 & 0 & 0 & 0 & c & h \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & a \end{array}\right].$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

Let S be a basic Weyr matrix of Weyr structure $(k_1, k_2, \ldots, k_r)$. Whenever we write $X \in \mathscr{C}(W)$ by its first row of blocks,

$$X = \left[\begin{array}{c|c|c|c} X_{1,1} & X_{1,2} & \cdots & X_{1,r} \end{array}\right], \tag{2.3}$$

we mean

$$X = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,r} \\ O & X_{2,2} & \cdots & X_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & X_{r,r} \end{bmatrix}, \tag{2.4}$$

where $X_{i,j}$ satisfies (2.2) in Lemma 2.7 for all i, j. We call (2.3) the *simplest form* of X and (2.4) the *expanded form* of X. If S is basic Weyr matrix with homogeneous Weyr structure of length 3, the matrices that commute with S are of the form

$$X = \begin{bmatrix} X_1 & X_2 & X_3 \\ 0 & X_1 & X_2 \\ 0 & 0 & X_1 \end{bmatrix},$$

in expanded form and $X = \left[\begin{array}{c|c|c} X_1 & X_2 & X_3 \end{array}\right]$ in simplest form where X_i satisfies (2.2) in Lemma 2.7.

We now prove Theorem 1.1.

3. Proof of Theorem 1.1. Suppose that S is a basic Weyr matrix with Weyr structure $(1, 1, \ldots, 1)$ of length r. It follows from Lemma 2.7 that $X \in \mathscr{C}(S)$ if and only if X has a simplest form given by

$$X = \left[\begin{array}{c|c|c|c} a_1 & a_2 & \cdots & a_r \end{array}\right],$$

for some $a_i \in \mathbb{F}$, $i = 1, 2, \ldots, r$. If X is an involution, then $X^2 = I$ and so

$$X^2 = \left[\begin{array}{c|c|c|c} b_1 & b_2 & \cdots & b_r \end{array}\right] = \left[\begin{array}{c|c|c|c} 1 & 0 & \cdots & 0 \end{array}\right],$$

where

$$b_j = \sum_{i=1}^{j} a_i a_{j-i+1}.$$

If $j = 1$, then $b_1 = a_1^2 = 1$. Hence, $a_1 = \pm 1$. Thus,

$$X = \left[\begin{array}{c|c|c|c} 1 & a_2 & \cdots & a_r \end{array}\right] \quad \text{or} \quad X = \left[\begin{array}{c|c|c|c} -1 & a_2 & \cdots & a_r \end{array}\right].$$

Since $X^2 = I$, X is diagonalizable and so the preceding tells us that either $X = I$ or $X = -I$. Therefore, in general if S is a basic Weyr matrix with Weyr structure $(1, 1, \ldots, 1)$ of length r, a matrix $X \in \mathscr{C}(S)$, $X \neq \pm I_r$ is not a product of involutions in $\mathscr{C}(S)$.

Let S be a basic Weyr matrix with homogeneous Weyr structure $(k, k, \ldots, k)$ of length $r \leq 3$. If $r = 1$, then $S = \lambda I$ for some eigenvalue $\lambda \in \mathbb{F}$ and every matrix in $\mathscr{C}(S)$ with determinant ± 1 is a product of involutions in $\mathscr{C}(S)$. Hence, we let $k \geq 2$ and $r \in \{2, 3\}$. For $A \in M_n(\mathbb{F})$, denote $\mathrm{tr}(A)$ the trace of A.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

THEOREM 3.1. *Let S be a basic Weyr matrix with Weyr structure $(k,k,\ldots,k)$ of length $r\geq 2$ where $k\geq 2$. If $X=\left[\begin{array}{c|c|c|c} X_{1,1} & X_{1,2} & \cdots & X_{1,r}\end{array}\right]$ is an involution in $\mathscr{C}(S)$, then $X_{1,1}$ is an involution,* $\operatorname{tr}(X_{1,1}X_{1,2})=0$, *and* $2\operatorname{tr}(X_{1,1}X_{1,3})=\operatorname{tr}((X_{1,1}X_{1,2})^2)$.

Proof. If $X=\left[\begin{array}{c|c|c|c} X_{1,1} & X_{1,2} & \cdots & X_{1,r}\end{array}\right]$ is an involution, then

$$X^2=\left[\begin{array}{c|c|c|c|c|c} A_1 & A_2 & A_3 & * & \cdots & *\end{array}\right]=I.$$

where $A_1=X_{1,1}^2$, $A_2=X_{1,1}X_{1,2}+X_{1,2}X_{1,1}$, and $A_3=X_{1,1}X_{1,3}+X_{1,2}^2+X_{1,3}X_{1,1}$. Thus, $X_{1,1}^2=I_k$ and $X_{1,1}X_{1,2}+X_{1,2}X_{1,1}=0$, and $X_{1,1}X_{1,3}+X_{1,2}^2+X_{1,3}X_{1,1}=0$. We get that

$$\begin{aligned}0&=\operatorname{tr}(X_{1,1}X_{1,2}+X_{1,2}X_{1,1})\\&=\operatorname{tr}(X_{1,1}X_{1,2})+\operatorname{tr}(X_{1,2}X_{1,1})\\&=2\operatorname{tr}(X_{1,1}X_{1,2}),\end{aligned}$$

and so $\operatorname{tr}(X_{1,1}X_{1,2})=0$. Also,

$$\begin{aligned}0&=\operatorname{tr}(X_{1,1}X_{1,3}+X_{1,2}^2+X_{1,3}X_{1,1})\\&=\operatorname{tr}(X_{1,1}X_{1,3})+\operatorname{tr}(X_{1,2}^2)+\operatorname{tr}(X_{1,3}X_{1,1})\\&=2\operatorname{tr}(X_{1,1}X_{1,3})+\operatorname{tr}(X_{1,2}^2),\end{aligned}$$

which implies that

$$2\operatorname{tr}(X_{1,1}X_{1,3})=-\operatorname{tr}(X_{1,2}^2). \tag{3.5}$$

Since $X_{1,1}X_{1,2}+X_{1,2}X_{1,1}=0$, then $X_{1,1}X_{1,2}=-X_{1,2}X_{1,1}$. Thus, we have that

$$\begin{aligned}\operatorname{tr}((X_{1,1}X_{1,2})^2)&=\operatorname{tr}(X_{1,1}X_{1,2}X_{1,1}X_{1,2})\\&=\operatorname{tr}(-X_{1,2}X_{1,1}X_{1,1}X_{1,2})\\&=-\operatorname{tr}(X_{1,2}^2).\end{aligned} \tag{3.6}$$

From (3.5) and (3.6), we obtain $2\operatorname{tr}(X_{1,1}X_{1,3})=\operatorname{tr}((X_{1,1}X_{1,2})^2)$. ◻

THEOREM 3.2. *Let S be a basic Weyr matrix with Weyr structure $(k,k,\ldots,k)$ of length $r\geq 2$ where $k\geq 2$ and let $X\in\mathscr{C}(S)$. If X satisfies the condition in Theorem 3.1, then so does X^{-1}.*

Proof. Suppose $X=\left[\begin{array}{c|c|c|c|c} X_{1,1} & X_{1,2} & X_{1,3} & \cdots & X_{1,r}\end{array}\right]$ such that $X_{1,1}$ is an involution, $\operatorname{tr}(X_{1,1}X_{1,2})=0$, and $2\operatorname{tr}(X_{1,1}X_{1,3})=\operatorname{tr}((X_{1,1}X_{1,2})^2)$. Then X is nonsingular and

$$X^{-1}=\left[\begin{array}{c|c|c|c|c|c} A_1 & A_2 & A_3 & * & \cdots & *\end{array}\right]\in\mathscr{C}(S).$$

where

$$A_1=X_{1,1}^{-1},\quad A_2=-X_{1,1}^{-1}X_{1,2}X_{1,1}^{-1},\quad A_3=X_{1,1}^{-1}X_{1,2}X_{1,1}^{-1}X_{1,2}X_{1,1}^{-1}-X_{1,1}^{-1}X_{1,3}X_{1,1}^{-1}.$$

Note that $A_1^{-1}=(X_{1,1}^{-1})^{-1}=X_{1,1}=X_{1,1}^{-1}=A_1$. Furthermore,

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

$$\begin{aligned}\operatorname{tr}(A_1A_2) &= \operatorname{tr}\left(X_{1,1}^{-1}(-X_{1,1}^{-1}X_{1,2}X_{1,1}^{-1})\right)\\ &= \operatorname{tr}\left(-(X_{1,1}^{-1})^2X_{1,2}X_{1,1}^{-1}\right)\\ &= \operatorname{tr}\left(-(X_{1,1})^2X_{1,2}X_{1,1}\right)\\ &= \operatorname{tr}(-X_{1,2}X_{1,1})\\ &= -\operatorname{tr}(X_{1,1}X_{1,2})\\ &= 0.\end{aligned}$$

by the assumption on X which proves that the second condition holds for X^{-1}. Finally,

$$\begin{aligned}\operatorname{tr}(A_1A_3) &= \operatorname{tr}(X_{1,2}X_{1,1}X_{1,2}X_{1,1} - X_{1,3}X_{1,1})\\ &= \operatorname{tr}(X_{1,1}X_{1,2}X_{1,1}X_{1,2}) - \operatorname{tr}(X_{1,1}X_{1,3})\\ &= \operatorname{tr}\left((X_{1,1}X_{1,2})^2\right) - \operatorname{tr}(X_{1,1}X_{1,3})\\ &= \frac{1}{2}\operatorname{tr}\left((X_{1,1}X_{1,2})^2\right).\end{aligned} \tag{3.7}$$

Now,

$$\begin{aligned}\operatorname{tr}\left((A_1A_2)^2\right) &= \operatorname{tr}\left((-X_{1,2}X_{1,1})^2\right)\\ &= \operatorname{tr}\left((X_{1,1}X_{1,2})^2\right).\end{aligned} \tag{3.8}$$

From (3.7) and (3.8), and the assumption on X that the third condition also holds for X^{-1}. ∎

The properties of X in Theorem 3.1 motivate us to study the following properties of a matrix in $\mathscr{C}(S)$. Let S be a basic Weyr matrix with Weyr structure $(k,k,k\ldots,k)$ where $k \geq 2$ of length $r \geq 2$. By Lemma 2.7, the simplest form of a matrix X in $\mathscr{C}(S)$ is of the form

$$X = \left[\begin{array}{c|c|c|c|c} X_{1,1} & X_{1,2} & X_{1,3} & \cdots & X_{1,r}\end{array}\right].$$

We say that X has property

- (P1) if $\det X_{1,1} = \pm 1$;
- (P2) if $\operatorname{tr}(X_{1,1}^{-1}X_{1,2}) = 0$; and
- (P3) if $2\operatorname{tr}(X_{1,1}^{-1}X_{1,3}) = \operatorname{tr}((X_{1,1}^{-1}X_{1,2})^2)$.

Consider the matrices $X_1 = \left[\begin{array}{c|c|c|c|c} A_1 & A_2 & A_3 & \cdots & A_r\end{array}\right]$ and $X_2 = \left[\begin{array}{c|c|c|c|c} B_1 & B_2 & B_3 & \cdots & B_r\end{array}\right]$ in $\mathscr{C}(S)$. Assume that X_1 and X_2 satisfy (P1), (P2), and (P3). We have

$$X_1X_2 = \left[\begin{array}{c|c|c|c|c} C_1 & C_2 & C_3 & \cdots & C_r\end{array}\right],$$

where

$$C_1 = A_1B_1, \quad C_2 = A_1B_2 + B_2A_1, \quad C_3 = A_1B_3 + A_2B_2 + A_3B_1.$$

Note that

$$\det C_1 = (\det A_1)(\det B_1) = (\pm 1)(\pm 1) = \pm 1.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

Hence (P1) holds for X_1X_2. Furthermore,

$$\begin{aligned}\operatorname{tr}(C_1^{-1}C_2) &= \operatorname{tr}\left((A_1B_1)^{-1}(A_1B_2 + A_2B_1)\right)\\ &= \operatorname{tr}(B_1^{-1}A_1^{-1}A_1B_2 + B_1^{-1}A_1^{-1}A_2B_1)\\ &= \operatorname{tr}(B_1^{-1}B_2) + \operatorname{tr}(A_1^{-1}A_2)\\ &= 0 + 0 = 0,\end{aligned}$$

by the assumption that X_1 and X_2 have property (P2). Hence (P2) holds for X_1X_2. Now, observe that

$$\begin{aligned}\operatorname{tr}(C_1^{-1}C_3) &= \operatorname{tr}\left((A_1B_1)^{-1})(A_1B_3 + A_2B_2 + A_3B_1)\right)\\ &= \operatorname{tr}\left((B_1^{-1}A_1^{-1})(A_1B_3 + A_2B_2 + A_3B_1)\right)\\ &= \operatorname{tr}(B_1^{-1}A_1^{-1}A_1B_3 + B_1^{-1}A_1^{-1}A_2B_2 + B_1^{-1}A_1^{-1}A_3B_1)\\ &= \operatorname{tr}(B_1^{-1}B_3) + \operatorname{tr}(B_1^{-1}A_1^{-1}A_2B_2) + \operatorname{tr}(A_1^{-1}A_3).\end{aligned} \tag{3.9}$$

Notice that

$$\begin{aligned}C_1^{-1}C_2 &= (B_1^{-1}A_1^{-1})(A_1B_2 + A_2B_1)\\ &= B_1^{-1}A_1^{-1}A_1B_2 + B_1^{-1}A_1^{-1}A_2B_1\\ &= B_1^{-1}B_2 + B_1^{-1}A_1^{-1}A_2B_1.\end{aligned}$$

Thus,

$$\begin{aligned}\operatorname{tr}\left((C_1^{-1}C_2)^2\right) &= \operatorname{tr}\left((B_1^{-1}B_2)^2\right) + 2\operatorname{tr}(B_1^{-1}A_1^{-1}A_2B_2) + \operatorname{tr}\left((A_1^{-1}A_2)^2\right)\\ &= 2\operatorname{tr}(B_1^{-1}B_3) + 2\operatorname{tr}(B_2B_1^{-1}A_1^{-1}A_2) + 2\operatorname{tr}(A_1^{-1}A_3).\end{aligned} \tag{3.10}$$

From (3.9) and (3.10), we have that $2\operatorname{tr}(C_1^{-1}C_3) = \operatorname{tr}\left((C_1^{-1}C_2)^2\right)$. Thus (P3) holds for X_1X_2 and so the product of matrices that satisfy the three properties also satisfies the same properties. One can inductively show that if $X_i \in \mathscr{C}(S)$ satisfies the three properties for all $i \in I$, then $X = \prod_{i\in I} X_i$ also satisfies (P1), (P2) and (P3).

THEOREM 3.3. *Let S be a basic Weyr matrix with Weyr structure $(k, k, \ldots, k)$ of length $r \geq 2$ where $k \geq 2$. Suppose $X_i \in \mathscr{C}(S)$ satisfies (P1), (P2), and (P3) for all $i = 1, 2, \ldots, s$. Then $X = \prod_{i=1}^{s} X_i$ also satisfies (P1), (P2), and (P3).*

The preceding discussion and Theorem 3.1 proves the forward implications of Theorems 1.1(a) and 1.1(b). We are left to prove its converse. Before we do this, we first prove several lemmas.

Denote $E_{i,j}$ the matrix such that its (i, j)-th entry is 1 and the rest of the entries are 0 and we define a matrix $D_j^{(n)}$ to be an n-by-n diagonal matrix whose entries are -1 in the j^{th} position of the diagonal and 1 in the rest of the diagonal entries. For example, $D_3^{(7)} = \operatorname{diag}(1, 1, -1, 1, 1, 1, 1)$.

LEMMA 3.4. *Let S be a basic Weyr matrix with Weyr structure $(k, k, \ldots, k)$ of length $r \geq 2$ where $k \geq 2$. For a fixed $s \in \{1, 2, \ldots, r-1\}$ and $i, j \in \{1, 2, \ldots, k\}$ such that $i \neq j$, let*

$$X = \left[\begin{array}{c|c|c|c|c} I_k & Y_1 & Y_2 & \cdots & Y_{r-1}\end{array}\right],$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

where

$$Y_m = \begin{cases} aE_{i,j} & \text{if } m = s \\ O_k & \text{otherwise} \end{cases},$$

for some $a \in \mathbb{F}$. Then X is a product of involutions in $\mathscr{C}(S)$.

Proof. Define $A = \left[\begin{array}{c|c|c|c|c} D_i^{(k)} & Y_1 & Y_2 & \cdots & Y_{r-1} \end{array}\right]$ where

$$Y_m = \begin{cases} aE_{i,j} & \text{if } m = s \\ O_k & \text{otherwise} \end{cases}.$$

Note that

$$A^2 = \left[\begin{array}{c|c|c|c|c} I_k & X_1 & X_2 & \cdots & X_{r-1} \end{array}\right],$$

where

$$X_m = D_i^{(k)} Y_m + Y_m D_i^{(k)} + \sum_{t=1}^{m-1} Y_t Y_{m-t}.$$

If $m < s$, then $X_m = O_k$ since $Y_t = O_k$ for all $t < s$. If $m > s$, then $X_m = O_k$ since $Y_t = O_k$ for all $t > s$. If $m = s$, then

$$\begin{aligned} X_s &= D_i^{(k)} Y_s + Y_s D_i^{(k)} \\ &= D_i^{(k)} (aE_{i,j}) + (aE_{i,j})(D_i^{(k)}) \\ &= -aE_{i,j} + aE_{i,j} = O_k, \end{aligned}$$

since $Y_t = O_k$ for all $t \neq s$. Hence, $X_m = O_k$ for all $m = 1, 2, \ldots, r-1$. Thus, $A^2 = I$ which implies that A is an involution in $\mathscr{C}(S)$.

Define $B = \bigoplus_{t=1}^{r} D_i^{(k)} \in \mathscr{C}(S)$. Then B is an involution since each $D_i^{(k)}$ is an involution. Now,

$$AB = \left[\begin{array}{c|c|c|c|c} D_i^{(k)} D_i^{(k)} & Y_1 D_i^{(k)} & Y_2 D_i^{(k)} & \cdots & Y_{r-1} D_i^{(k)} \end{array}\right].$$

Since $Y_i = O_k$ for all $i \neq s$, then

$$\begin{aligned} AB &= \left[\begin{array}{c|c|c|c|c|c|c} I_k & O_k & O_k & \cdots & (aE_{i,j})(D_i^{(k)}) & \cdots & O_k \end{array}\right] \\ &= \left[\begin{array}{c|c|c|c|c|c|c} I_k & O_k & O_k & \cdots & aE_{i,j} & \cdots & O_k \end{array}\right] \\ &= X. \end{aligned}$$

Therefore, X is a product of involutions in $\mathscr{C}(S)$. ∎

Lemma 3.5. *Let S be a basic Weyr matrix with Weyr structure $(k, k, \ldots, k)$ of length $r \geq 2$ where $k \geq 2$. Let $A \in M_k(\mathbb{F})$ such that $\operatorname{tr}(A) = 0$. Then*

$$X = \left[\begin{array}{c|c|c|c|c} I_k & O_k & \cdots & O_k & A \end{array}\right],$$

is a product of involutions in $\mathscr{C}(S)$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

Proof. Let $A = [a_{i,j}] \in M_k(\mathbb{F})$ such that $\operatorname{tr}(A) = 0$, that is $\sum_{i=1}^{k} a_{i,i} = 0$. Let $a_i = \sum_{j=1}^{i} a_{j,j}$. Consider the matrices

$$X_i = \left[\begin{array}{c|c|c|c|c} I_k & O_k & \cdots & O_k & a_i(E_{i+1,i+1} - E_{i,i}) \end{array}\right],$$

for all $i = 1, 2, \ldots, k-1$. Note that $X_i = X_i^{(1)} X_i^{(2)}$ where

$$X_i^{(1)} = \left[\begin{array}{c|c|c|c|c} I_{i-1} \oplus \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] \oplus I_{k-i-1} & O_k & \cdots & O_k & a_i(E_{i+1,i} - E_{i,i+1}) \end{array}\right],$$

and

$$X_i^{(2)} = \left[\begin{array}{c|c|c|c|c} I_{i-1} \oplus \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] \oplus I_{k-i-1} & O_k & \cdots & O_k & O_k \end{array}\right].$$

One checks that $X_i^{(2)}$ is an involution in $\mathscr{C}(S)$ for all i. Furthermore,

$$\left(X_i^{(1)}\right)^2 = \left[\begin{array}{c|c|c|c|c} I_k & O_k & \cdots & O_k & O_{i-1} \oplus C \oplus O_{k-i-1} \end{array}\right],$$

where

$$\begin{aligned} C &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -a_i \\ a_i & 0 \end{bmatrix} + \begin{bmatrix} 0 & -a_i \\ a_i & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a_i & 0 \\ 0 & -a_i \end{bmatrix} + \begin{bmatrix} -a_i & 0 \\ 0 & a_i \end{bmatrix} \\ &= O_2. \end{aligned}$$

Hence, $X_i^{(1)}$ is an involution in $\mathscr{C}(S)$. Thus, each X_i is a product of involutions in $\mathscr{C}(S)$. Let $Y = X_{k-1}X_{k-2}\cdots X_2X_1X$. Observe that

$$Y = \left[\begin{array}{c|c|c|c|c} I_k & O_k & \cdots & O_k & B \end{array}\right],$$

where $B = [b_{i,j}] \in M_k(\mathbb{F})$ such that $b_{i,i} = 0$ for all $i = 1, 2, \ldots, k-1$ and $b_{k,k} = \sum_{i=1}^{k} a_{i,i}$. Since $\sum_{i=1}^{k} a_{i,i} = 0$, then $b_{k,k} = 0$, and by Lemma 3.4, Y is a product of involutions in $\mathscr{C}(S)$. Furthermore, since the X_i's are products of involutions in $\mathscr{C}(S)$, then so is X_i^{-1} by Theorem 3.2. Hence, $X = X_1^{-1}X_2^{-1}\cdots X_{k-1}^{-1}Y$ is a product of involutions in $\mathscr{C}(S)$. □

Let R_n be the n-by-n backward identity matrix, that is

$$R_n = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

We proceed with the following lemma.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

LEMMA 3.6. *Let S be a basic Weyr matrix with Weyr structure (k,k,k) where $k \geq 2$. Let $X = \left[\begin{array}{c|c|c} I_k & A & B \end{array}\right] \in \mathscr{C}(S)$ such that $\operatorname{tr}(A) = 0$, and $2\operatorname{tr}(B) = \operatorname{tr}(A^2)$. Then X is a product of involutions in $\mathscr{C}(S)$.*

Proof. Suppose $A = [a_{i,j}] \in M_k(\mathbb{F})$ and $B \in M_k(\mathbb{F})$ such that $\operatorname{tr}(A) = 0$ and $2\operatorname{tr}(B) = \operatorname{tr}(A^2)$. This means that X satisfies (P2) and (P3). Let $a_i = \sum_{j=1}^{i} a_{j,j}$ and consider the matrices

$$X_i = \left[\begin{array}{c|c|c} I_k & a_i(E_{i+1,i+1} - E_{i,i}) & a_i^2(E_{i+1,i+1}) \end{array}\right],$$

$$Y_i = \left[\begin{array}{c|c|c} I_{i-1} \oplus R_2 \oplus I_{k-i-1} & O_{i-1} \oplus \left[\begin{smallmatrix} 0 & -a_i \\ a_i & 0 \end{smallmatrix}\right] \oplus O_{k-i-1} & a_i^2 E_{i+1,i} \end{array}\right],$$

and

$$Z_i = \left[\begin{array}{c|c|c} I_{i-1} \oplus R_2 \oplus I_{k-i-1} & O_k & O_k \end{array}\right].$$

for all $i = 1, 2, \ldots, k-1$. One checks that Y_i and Z_i are involutions in $\mathscr{C}(S)$. Thus,

$$\begin{aligned} Y_i Z_i &= \left[\begin{array}{c|c|c} I_k & O_{i-1} \oplus \left[\begin{smallmatrix} -a_i & 0 \\ 0 & a_i \end{smallmatrix}\right] \oplus O_{k-i-1} & (a_i)^2(E_{i+1,i+1}) \end{array}\right] \\ &= \left[\begin{array}{c|c|c} I_k & a_i(E_{i+1,i+1} - E_{i,i}) & a_i^2(E_{i+1,i+1}) \end{array}\right] \\ &= X_i. \end{aligned}$$

Hence X_i is a product of involutions in $\mathscr{C}(S)$ and so

$$\begin{aligned} X_{k-1}X_{k-2}\cdots X_2X_1X &= \left[\begin{array}{c|c|c} I_k & A + \left(\sum_{i=1}^{k-1} a_i\right)(E_{i+1,i+1} - E_{i,i}) & B_1 \end{array}\right] \\ &= \left[\begin{array}{c|c|c} I_k & A_1 & B_1 \end{array}\right], \end{aligned} \tag{3.11}$$

where

$$A_1 = \begin{bmatrix} 0 & a_{1,2} & \cdots & a_{1,k} \\ a_{2,1} & 0 & \cdots & a_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k,1} & a_{k,2} & \vdots & \left(\sum_{j=1}^{k-1} a_{j,j}\right) + a_{k,k} \end{bmatrix},$$

for some $B_1 \in M_k(\mathbb{F})$. Since $\operatorname{tr}(A) = 0$, then $\left(\sum_{j=1}^{k-1} a_{j,j}\right) + a_{k,k} = 0$. Let $Y = \left[\begin{array}{c|c|c} I_k & A_1 & B_1 \end{array}\right]$. Since X and the X_i's satisfy (P1), (P2), and (P3), then so does Y by Theorem 3.3. Now, for all $i, j = 1, 2, \ldots, k$ such that $i \neq j$, consider the matrices $Y_{i,j}$ given by

$$Y_{i,j} = \left[\begin{array}{c|c|c} I_k & -a_{i,j}E_{i,j} & O_k \end{array}\right].$$

By Lemma 3.4, $Y_{i,j}$ is a product of involutions in $\mathscr{C}(S)$ for all $i \neq j$ and hence

$$\begin{aligned} Y\left(\prod_{i\neq j} Y_{i,j}\right) &= \left[\begin{array}{c|c|c} I_k & A_1 - \sum_{i\neq j} a_{i,j}E_{i,j} & B_2 \end{array}\right] \\ &= \left[\begin{array}{c|c|c} I_k & O_k & B_2 \end{array}\right]. \end{aligned} \tag{3.12}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

for some $B_2 \in M_k(\mathbb{F})$. Since Y and $Y_{i,j}$ satisfy (P1), (P2), and (P3), then so does the matrix $Z := \left[\ I_k \ \middle|\ O_k \ \middle|\ B_2\ \right]$ by Theorem 3.3 which implies that $\mathrm{tr}(B_2) = 0$ since $2\mathrm{tr}(B_2) = \mathrm{tr}(O_k^2) = 0$. By Lemma 3.5, Z is a product of involutions in $\mathscr{C}(S)$. From Equations (3.11) and (3.12), we have that X is a product of involutions in $\mathscr{C}(S)$. ◻

We now prove the backward implication of Theorem 1.1.

Proof of Theorem 1.1. We only prove Theorem 1.1 b) since the proof of Theorem 1.1 a) follows similarly. Assume that $\det X_1 = \pm 1$, $\mathrm{tr}(X_1^{-1}X_2) = 0$ and $2\mathrm{tr}(X_1^{-1}X_3) = \mathrm{tr}((X_1^{-1}X_2)^2)$. Note that

$$X = \left[\ X_1 \ \middle|\ O_k \ \middle|\ O_k\ \right] \left[\ I_k \ \middle|\ X_1^{-1}X_2 \ \middle|\ X_1^{-1}X_3\ \right].$$

and the fact that $\det X_1 = \pm 1$ implies that X_1 is a product of involutions and we can assume that $X_{1,1}$ is a product of four involutions, say $X_1 = X_{1,1}X_{2,2}X_{3,3}X_{4,4}$ where $X_{i,i}^2 = I$ by [1]. Take $Y_i = X_{i,i} \oplus X_{i,i} \oplus X_{i,i}$ for all $i = 1, 2, 3, 4$. Then Y_i is an involution in $\mathscr{C}(S)$ and so $\left[\ X_1 \ \middle|\ O_k \ \middle|\ O_k\ \right] = Y_1Y_2Y_3Y_4$ is a product of involutions in $\mathscr{C}(S)$. Let

$$Y = \left[\ I_k \ \middle|\ X_1^{-1}X_2 \ \middle|\ X_1^{-1}X_3\ \right].$$

Since $\mathrm{tr}(X_1^{-1}X_2) = 0$ and $2\mathrm{tr}(X_1^{-1}X_3) = \mathrm{tr}((X_1^{-1}X_2)^2)$, then Y is a product of involutions in $\mathscr{C}(S)$ by Lemma 3.6. Therefore, X is a product of involutions in $\mathscr{C}(S)$. ◻

We investigate the case when the length of the Weyr structure is greater than 3.

4. Extended results.

Assume S is a basic $n \times n$ Weyr matrix of homogeneous structure $(k, k, \ldots, k)$ with r lots of k (so $n = kr$).

Theorem 4.1. *Every invertible matrix $X \in \mathscr{C}(S)$ can be written as a product $X = X_1X_2 \cdots X_r$ of matrices $X_i \in \mathscr{C}(S)$ whose top row (simplest) forms are*

$$X_i = \left[\ X_{i1} \ \middle|\ 0 \ \middle|\ \cdots \ \middle|\ X_{ii} \ \middle|\ 0 \ \middle|\ \cdots \ \middle|\ 0\ \right],$$

with an invertible 1st component and 0's in all other components except the ith.

Proof. We illustrate the case when $r = 4$ and $X = [\ A \mid B \mid C \mid D\ ]$, that is in expanded form

$$X = \begin{bmatrix} A & B & C & D \\ & A & B & C \\ & & A & B \\ & & & A \end{bmatrix}.$$

Since $X \in M_n(F)$ is invertible and block upper triangular, $A \in M_k(F)$ must be invertible. By factoring out the diagonal, we can assume $A = I_k$. Using essentially elementary column operations but in block form, we can reduce X to I_n as follows:

$$\begin{bmatrix} I & B & C & D \\ & I & B & C \\ & & I & B \\ & & & I \end{bmatrix} \begin{bmatrix} I & -B & 0 & 0 \\ & I & -B & 0 \\ & & I & -B \\ & & & I \end{bmatrix} = \begin{bmatrix} I & 0 & -B^2 + C & -CB + D \\ & I & 0 & -B^2 + C \\ & & I & 0 \\ & & & I \end{bmatrix}.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

$$\begin{bmatrix} I & 0 & -B^2+C & -CB+D \\ & I & 0 & -B^2+C \\ & & I & 0 \\ & & & I \end{bmatrix} \begin{bmatrix} I & 0 & B^2-C & 0 \\ & I & 0 & B^2-C \\ & & I & 0 \\ & & & I \end{bmatrix} = \begin{bmatrix} I & 0 & 0 & -CB+D \\ & I & 0 & 0 \\ & & I & 0 \\ & & & I \end{bmatrix}.$$

$$\begin{bmatrix} I & 0 & 0 & -CB+D \\ & I & 0 & 0 \\ & & I & 0 \\ & & & I \end{bmatrix} \begin{bmatrix} I & 0 & 0 & CB-D \\ & I & 0 & 0 \\ & & I & 0 \\ & & & I \end{bmatrix} = \begin{bmatrix} I & 0 & 0 & 0 \\ & I & 0 & 0 \\ & & I & 0 \\ & & & I \end{bmatrix}.$$

Hence

$$X^{-1} = \begin{bmatrix} I & -B & 0 & 0 \\ & I & -B & 0 \\ & & I & -B \\ & & & I \end{bmatrix} \begin{bmatrix} I & 0 & B^2-C & 0 \\ & I & 0 & B^2-C \\ & & I & 0 \\ & & & I \end{bmatrix} \begin{bmatrix} I & 0 & 0 & CB-D \\ & I & 0 & 0 \\ & & I & 0 \\ & & & I \end{bmatrix}.$$

The last product of 3 matrices has the required form for the invertible matrix $X^{-1} \in \mathscr{C}(S)$. We get the form for X $(= (X^{-1})^{-1})$ by starting with X^{-1} in simplest form instead of X. $\square$

The problem of characterizing invertible matrices $X \in \mathscr{C}(S)$ that are expressible as products of involutions in $\mathscr{C}(S)$ has been "largely reduced" to considering matrices of the form

$$X = \begin{bmatrix} A & 0 & \cdots & 0 & Y & 0 & \cdots & 0 \\ & A & & & & Y & & \\ & & A & & & & Y & \\ & & & \ddots & & & & \ddots \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & A \end{bmatrix},$$

where $\det(A) = \pm 1$ and $Y \in M_k(F)$ is arbitrary. Moreover, we can assume $A = I$ by Theorem 4.1. We say it has been "largely reduced" because although it would be sufficient to have each of the terms X_i in the factorization $X = X_1 X_2 \cdots X_r$ a product of involutions to deduce the same for X, it is not clear from a factorization such as given in the proof of Theorem 4.1 that the converse holds.

When X in simplest form is $[\ I \mid B\ ]$ (case $r = 2$), X is a product of involutions exactly when $\mathrm{tr}(B) = 0$ by Theorem 1.1. For $r = 3$ and $X = [\ I \mid B \mid C\ ]$, the conditions are $\mathrm{tr}(B) = 0$ and $\mathrm{tr}(B^2) = 2\mathrm{tr}(C)$ by Theorem 1.1. It appears that further trace conditions on various products and powers of the Weyr structure components are required for larger r.

Theorem 4.2. *Let $X = [\ I \mid \cdots \mid 0 \mid Y \mid 0 \mid \cdots\ ]$ where Y is the mth component and $m \geq 1 + r/2$. If $\mathrm{tr}(Y) = 0$, then X is a product of involutions in $\mathscr{C}(S)$. In characteristic zero, if $m \leq 1 + r/2$ then $\mathrm{tr}(Y) = 0$ is necessary for X to be a product of involutions.*

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

Proof. Fix m with $1+r/2 \le m \le r$. The map

$$\begin{bmatrix} A & Y \\ & A \end{bmatrix} \quad \longmapsto \quad \begin{bmatrix} A & 0 & \cdots & 0 & Y & 0 & \cdots & 0 \\ & A & & & & Y & & \\ & & A & & & & Y & \\ & & & \ddots & & & & \ddots \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & A \end{bmatrix},$$

where the Y 's in the image occupy the $m-1$ st superdiagonal, preserves multiplication and the identity because of our restriction on m. Thus products of involutions in the centralizer of the Weyr matrix of structure (k, k) (the case $r = 2$) are mapped to products of involutions in $\mathscr{C}(S)$. We know $\begin{bmatrix} I & Y \\ & I \end{bmatrix}$ is a product of involutions when $\mathrm{tr}(Y) = 0$. Therefore, having $\mathrm{tr}(Y) = 0$ in the image matrix is sufficient for the matrix X to be a product of involutions in $\mathscr{C}(S)$.

Now suppose we consider a field of characteristic zero and $1 < m \le 1 + r/2$. Since $2m-2 \le r$, we have the projection map π of $\mathscr{C}(S)$ onto the top left $(2m-2)\times(2m-2)$ corner of $k \times k$ blocks, which is a ring homomorphism and preserves the identity. Therefore, π maps products of involutions in $\mathscr{C}(S)$ to products of involutions. Let $T = \pi(\mathscr{C}(S))$, which is naturally isomorphic to the centralizer of the Weyr matrix with structure $(k, k, \ldots, k)$, $2m-2$ lots of k. Next block T as a 2×2 block matrix ring. Then T can be viewed as sitting inside $\mathscr{C}(V)$ where V is the Weyr matrix of structure $(k(m-1), k(m-1))$ (a case $r = 2$). Viewed as a member of $\mathscr{C}(V)$, we have

$$\pi(X) = \begin{bmatrix} I & \bar{Y} \\ 0 & I \end{bmatrix},$$

where $\bar{Y} = \mathrm{diag}(Y, Y, \ldots, Y)$, $m-1$ copies of Y. Hence if X is a product of involutions in $\mathscr{C}(S)$, then $\pi(X) = \begin{bmatrix} I & \bar{Y} \\ 0 & I \end{bmatrix}$ is a product of involutions in $\mathscr{C}(V)$. By the characterization for the case $r = 2$, we have $0 = \mathrm{tr}(\bar{Y}) = (m-1)\mathrm{tr}(Y)$, which implies $\mathrm{tr}(Y) = 0$ in characteristic 0. ∎

THEOREM 4.3. *If $X = [\, I \mid \cdots \mid 0 \mid Y \mid 0 \mid \cdots \,]$ where Y is nilpotent, then X is a product of involutions in $\mathcal{C}(S)$.*

Proof. Y is similar to a strictly upper triangular matrix, say $Q^{-1}YQ$ where $Q \in M_k(F)$ is invertible. Conjugation by

$$P = \mathrm{diag}(Q, Q, \ldots, Q), \quad r \text{ copies of } Q,$$

is an automorphism of $\mathscr{C}(S)$ and $P^{-1}XP = [\, I \mid \cdots \mid 0 \mid Q^{-1}YQ \mid 0 \mid \cdots \,]$. Thus, we can assume $Y = [y_{ij}]$ is strictly upper triangular. Left multiplying X by $Z = [\, I \mid \cdots \mid 0 \mid E \mid 0 \mid \cdots \,]$ with $E = -y_{ij}E_{ij}$ in the mth position and $i < j$ removes entry y_{ij} of Y. So a series of such multiplications makes Y zero. By Lemma 3.4, each Z is a product of involutions, whence

$$X = \text{product of involutions} \times [\, I \mid \cdots \mid 0 \mid Y \mid * \mid \cdots \mid * \,],$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

where now $Y=0$. If Y occupied the mth position and $m \leq 1+r/2$, we may have introduced a nonzero component in position $2m-1$. However, that component will be strictly upper triangular, so we can repeat the argument to eventually get

$$X \ = \ \text{product of involutions} \ \times \ [\ I \mid \cdots \mid 0 \mid 0 \mid \cdots \mid 0\ ].$$

Thus, X is a product of involutions. ∎

We have now largely reduced the product of involutions to matrices of the form

$$\begin{bmatrix} I & 0 & \cdots & 0 & D & 0 & \cdots & 0 \\ & I & & & & D & & \\ & & I & & & & D & \\ & & & \ddots & & & & \ddots \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & I \end{bmatrix},$$

where D is a diagonal matrix. This is because Y is similar to an upper triangular matrix and we can remove the strictly upper entries by Lemma 3.4.

Example 4.4. Let $r=4$. To characterize when

$$X \ = \ \begin{bmatrix} I & B & C & D \\ & I & B & C \\ & & I & B \\ & & & I \end{bmatrix},$$

is a product of involutions, for sufficiency it is enough to do this for X^{-1} using the factorization

$$X^{-1} \ = \ \begin{bmatrix} I & -B & & \\ & I & -B & \\ & & I & -B \\ & & & I \end{bmatrix} \begin{bmatrix} I & 0 & B^2-C & \\ & I & 0 & B^2-C \\ & & I & 0 \\ & & & I \end{bmatrix} \begin{bmatrix} I & 0 & 0 & CB-D \\ & I & 0 & 0 \\ & & I & 0 \\ & & & I \end{bmatrix},$$

in the proof of Theorem 4.1.

Necessary conditions. The projection of X to the 3×3 corner is a product of involutions, hence by case $r=3$ we have

1. $\mathrm{tr}(B)=0$.
2. $\mathrm{tr}(B^2)=2\mathrm{tr}(C)$.

Sufficient conditions in characteristic 0. Imposing the condition

3. $\mathrm{tr}(BC)=\mathrm{tr}(D)$

ensures the 3rd factor of X^{-1} is a product of involutions by Theorem 4.2, while imposing

4. $\mathrm{tr}(B^2)=\mathrm{tr}(D)$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

also ensures the 2nd factor is a product of involutions by Theorem 4.2. Requiring

5. $\operatorname{tr}(B^{i}) = 0$ for $i = 1, 2\ldots, k$

would make B nilpotent in characteristic 0 and hence by Theorem 4.3 the 1st factor of X^{-1} is a product of involutions. Thus, taken together (3), (4), and (5) are sufficient to give X^{-1}, and therefore X, as a product of involutions.

The Weyr form lends itself well to inductive arguments on the length r of the Weyr structure. For future work, a profitable line might be to use the observation that products of involutions for length r 'extend' to products of involutions of length $r+1$. That is:

THEOREM 4.5. *If S and T are Weyr matrices of structure $(k, k\ldots, k)$ and of lengths r and $r+1$ respectively, and $X = [\ X_{11} \mid X_{22} \mid X_{33} \mid \ \cdots \mid X_{r-1,r-1} \mid X_{rr}\ ]$ is a product of involutions in $\mathscr{C}(S)$, then there exists $X_{r+1,r+1} \in M_k(F)$ such that $X' = [\ X_{11} \mid X_{22} \mid X_{33} \mid \ \cdots \mid X_{r-1,r-1} \mid X_{rr} \mid X_{r+1,r+1}\ ]$ is a product of involutions of $\mathscr{C}(T)$.*

We state the following lemma before proving Theorem 4.5

LEMMA 4.6. *Let R be a ring (with identity) in which 2 is invertible, and let N be a nilpotent ideal of R. Then involutions of $\bar{R} = R/N$ 'lift' to involutions of R. That is, if $\bar{u} \in \bar{R}$ is an involution, then there exists an involution $v \in R$ such that $\bar{v} = \bar{u}$. (Here $\bar{u} = u + N \in R/N$ etc.)*

Proof. Since 2 is invertible, involutions u are precisely the elements of the form

$$u \ = \ 2e \ - \ 1, \quad \textit{where } e = e^2 \textit{ is an idempotent}.$$

Just note $(2e-1)^2 = 1$ and if $u^2 = 1$, then $e = (u+1)/2$ is an idempotent. Therefore, $\bar{u} = 2\bar{e} - 1$ for some idempotent $\bar{e} \in \bar{R}$. But it is well-known that idempotents lift modulo nilpotent ideals [4, Proposition 7.14]. So there is an idempotent $f \in R$ such that $\bar{f} = \bar{e}$. Let $v = 2f - 1$. Then v is an involution and $\bar{v} = 2\bar{f} - 1 = 2\bar{e} - 1 = \bar{u}$. ∎

Proof of Theorem 4.5. Those $X \in \mathscr{C}(T)$ whose top row forms are $X = [\ 0 \mid 0 \mid \ \cdots \mid 0 \mid X_{r+1,r+1}\ ]$ form a nilpotent ideal of $R = \mathscr{C}(T)$, in fact N is the ideal generated by T^r. Note $N = \ker(\pi)$ where $\pi : \mathscr{C}(T) \to \mathscr{C}(S)$ is the projection of $\mathscr{C}(T)$ onto its top left $r \times r$ corner of $k \times k$ blocks. Since π is a ring homomorphism, we have $\mathscr{C}(S) \cong R/N$. By Lemma 4.6, products of involutions of R/N lift to products of involutions of R.

Theorem 4.5 fails in characteristic 2: the involution

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

does not extend to an involution

$$\begin{bmatrix} 1 & 1 & z \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

because that would require $0 = 2z + 1 = 1$.

The proof given in Jacobson [4, Proposition 7.14] for lifting an idempotent $\bar{e}$ to an idempotent f is quite constructive if we know the nilpotent index of N. For the case we are interested in, N has nilpotent index

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

2 and the proof shows we can take

$$f \;=\; 3e^2 - 2e^3.$$

Retracing our proof, when the nilpotent ideal N has index 2, by taking $e = (1+u)/2$, we see an involution $\bar{u}$ lifts to the involution

$$\begin{aligned} v \;&=\; 2f - 1 \;= 6e^2 - 4e^3 - 1 \\ &=\; (3u - u^3)/2. \end{aligned}$$

This enables us to write down explicitly an extension of an involution $X \in \mathscr{C}(S)$ to one V of length $r+1$. For example, when $r = 3$ and

$$X \;=\; \begin{bmatrix} A & B & C \\ & A & B \\ & & A \end{bmatrix},$$

the matrix

$$U \;=\; \begin{bmatrix} A & B & C & 0 \\ & A & B & C \\ & & A & B \\ & & & A \end{bmatrix},$$

in $\mathscr{C}(T)$ is mapped to X under the projection π. Thus, by our lifting results, the involution $\pi(U) = \pi(V)$ where V is the involution

$$V \;=\; (3U - U^3)/2.$$

(Think of $\pi(X)$ as $\bar{U}$.) If we compute the right hand side, this is just saying the matrix

$$V \;=\; \begin{bmatrix} A & B & C & Z \\ & A & B & C \\ & & A & B \\ & & & A \end{bmatrix},$$

where $Z = -(ABC + ACB + BCA + CBA)/2$, is an involution which extends X. A similar calculation shows the matrix

$$\begin{bmatrix} A & B & -AB^2 \\ & A & B \\ & & A \end{bmatrix},$$

is an involution extending an involution

$$\begin{bmatrix} A & B \\ & A \end{bmatrix}.$$

Consider the basic Weyr matrix S with Weyr structure $(2,1,1,1)$ over a field of characteristic 2. Notice that

$$X = \left[\begin{array}{cc|c|c|c} 1 & a & 1 & z & e \\ 0 & b & c & d & d \\ \hline 0 & 0 & 1 & 1 & z \\ \hline 0 & 0 & 0 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right] \in \mathscr{C}(S),$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

is not an involution since the submatrix

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

of X is not an involution. Thus, in general, if S has homogeneous Weyr structure $(k, k, \ldots, k)$ of length r and T has Weyr structure $(k_1, k_2, \ldots, k, k, \ldots, k)$ and X is not an involution in $\mathscr{C}(S)$, then X does not extend to an involution in $\mathscr{C}(T)$.

5. Conclusions and future work. Let S be a basic Weyr matrix with homogeneous Weyr structure $(k, k, \ldots, k)$ of length r. Then $X \in \mathscr{C}(S)$ if and only if $X = \left[\begin{array}{c|c|c} X_1 & X_2 & X_3 \end{array}\right]$ in simplest form. If $r = 1$, then $S = \lambda I$ for some eigenvalue $\lambda \in \mathbb{F}$ and every matrix in $\mathscr{C}(S)$ with determinant ± 1 is a product of involutions in $\mathscr{C}(S)$. If $r = 2$, Theorem 1.1 a) guarantees that $\det X_1 = \pm 1$ and $\operatorname{tr}(X_1^{-1} X_2) = 0$ are both necessary and sufficient conditions for a matrix in $\mathscr{C}(S)$ to be a product of involutions in $\mathscr{C}(S)$. Finally, if $r = 3$, Theorem 1.1 b) states an added condition to the preceding, namely $2\operatorname{tr}(X_1^{-1} X_3) = \operatorname{tr}((X_1^{-1} X_2)^2)$, is needed for a matrix in $\mathscr{C}(S)$ to be a product of involutions in $\mathscr{C}(S)$. We conjecture that as the length of a basic Weyr matrix S increases, additional properties other than (P1), (P2), and (P3) are needed to conclude that a matrix is a product of involutions in $\mathscr{C}(S)$.

Theorem 4.5 suggests that in going from characterizing the length r case to characterizing length $r+1$, no new conditions are required that involve only the first r components, but rather only new conditions that involve the $r+1$ component and earlier components. We illustrate the case when $r = 3$ and characterizing when

$$X = \begin{bmatrix} A & B & C & D \\ & A & B & C \\ & & A & B \\ & & & A \end{bmatrix},$$

is a product of involutions of $\mathscr{C}(T)$. Projecting we know

$$\pi(X) = \begin{bmatrix} A & B & C \\ & A & B \\ & & A \end{bmatrix},$$

must be a product of involutions, and we know exactly the conditions for this to hold. Assuming these conditions, Theorem 4.5 guarantees a matrix

$$X' = \begin{bmatrix} A & B & C & Z \\ & A & B & C \\ & & A & B \\ & & & A \end{bmatrix},$$

which is a product of involutions. Therefore, X is a product of involutions if and only if the matrix $Y = X(X')^{-1}$ is such a product. But notice the special form of Y:

$$Y = \begin{bmatrix} I & 0 & 0 & E \\ & I & 0 & 0 \\ & & I & 0 \\ & & & I \end{bmatrix}.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 463-482, August 2022.

By Theorem 4.2, requiring $\mathrm{tr}(E) = 0$ is sufficient now for X to be a product of involutions. Is it also necessary?

Acknowledgment. The work of the first and second authors is supported by UP NSRI grant no. MAT-22-1-05. The authors express their gratitude to the anonymous referee who gave significant improvements to the original results of the authors.

REFERENCES

[1] W.H. Gustafson, P.R. Halmos, and H. Radjavi. Products of involutions. *Linear Algebra Appl.* 13:157–162, 1976.
[2] P. Lancaster and M. Tismenetsky. *The Theory of Matrices,* 2nd edn. Academic Press, Orlando, 1985.
[3] K.C. O'Meara, J. Clark, and C.I. Vinsonhaler, *Advanced Topics in Linear Algebra.* Oxford University Press, Oxford, 2011.
[4] N. Jacobson, *Basic Algebra II*, 2nd edn. W.H. Freeman and Company, New York, 1989.