SEMIGROUPS GENERATED BY SIMILARITY ORBITS

L. Grunenfelder, M. Omladič, H. Radjavi, A. Sourour

Abstract. We investigate the semigroups in $M_n(\mathbf{F})$ generated by the similarity orbit of single matrices.

0. Introduction

Question. What is the semigroup in $M_n(\mathbf{F})$ generated by the similarity orbit of a single matrix of rank k ?

In section 1 and 2 we consider the semigroup $\mathcal{S}$ in $M_n(\mathbf{F})$ generated by the similarity orbit of an invertible matrix A. In this case $\mathcal{S}$ is of course a semigroup in $\mathrm{GL}_n(\mathbf{F})$, and it is a normal subgroup if and only if $\det A$ is a root of unity in $\mathbf{F}^*$. For a non-scalar A, except when $n=2$ and $|\mathbf{F}| \leq 3$, these normal subgroups are isomorphic to semi-direct products $\mathcal{S} \cong \mathrm{SL}_n(\mathbf{F}) \bowtie U$, where U is the cyclic subgroup of $\mathbf{F}^*$ generated by $\det A$.

Some bounds for the number of similarity factors required are found in section 2. Let $(A)_m = \{A_1 A_2 \dots A_m | A_j \sim A \text{ for } j = 1, 2 \dots, m\}$. If $A = \lambda I$ is scalar, then of course $(A)_m$ is the singleton $\{\lambda^m I\}$. If A is not scalar, an obvious necessary condition for T to be in $(A)_m$ is that $\det(T) = (\det(A))^m$. We prove that this condition is sufficient if m is large enough. We find a bound on m in terms of the number of linear invariant factors of A; this bound never exceeds $4n$.

In section 3 we find that the semigroup in $M_n(\mathbf{F})$ generated by the similarity orbit of a singular matrix A with $\operatorname{rank} A = r < n$ consists of all matrices of rank less than or equal to r.

1. Semigroups generated by the similarity orbit of an invertible matrix

The semigroup $\mathcal{S}$ in $\mathrm{GL}_n(\mathbf{F})$ generated by the similarity orbit of a matrix A of finite multiplicative order is automatically a normal subgroup of $\mathrm{GL}_n(\mathbf{F})$. It is therefore useful to characterize the normal subgroups of $\mathrm{SL}_n(\mathbf{F})$ and of $\mathrm{GL}_n(\mathbf{F})$ first. Recall [AB] that $\mathrm{SL}_n(\mathbf{F})$ is perfect, i.e. $\mathrm{SL}_n(\mathbf{F})_{ab} = \mathrm{SL}_n(\mathbf{F})/[SL_n(\mathbf{F}), \mathrm{SL}_n(\mathbf{F})]$ is trivial, and

Research supported in part by the NSERC of Canada and by the Ministry of Science and Technology of Slovenia.

$\mathrm{PSL}_n(\mathbf{F})$ is simple when $n \neq 2$ and $|\mathbf{F}| \neq 2, 3$. Moreover, $Z(\mathrm{SL}_n(\mathbf{F})) = Z(\mathrm{GL}_n(\mathbf{F})) \cap \mathrm{SL}_n(\mathbf{F})$ for every field $\mathbf{F}$.

Lemma 1.1. *Let H be a normal subgroup of $\mathrm{SL}_n(\mathbf{F})$, where $n \neq 2$ and $|\mathbf{F}| \neq 2, 3$. Then either*

(1) *H consists of scalar matrices and is therefore a cyclic subgroup generated by an n-th root of unity, or*
(2) *H contains a non-scalar matrix and is equal to $\mathrm{SL}_n(\mathbf{F})$.*

Proof. If $Z = Z(\mathrm{SL}_n(\mathbf{F})) = \mathrm{SL}_n(\mathbf{F}) \cap Z(\mathrm{GL}_n(\mathbf{F}))$ is the center of $\mathrm{SL}_n(\mathbf{F})$, i.e. the cyclic subgroup of n-th roots of unity, then the obvious commutative diagram

$$\begin{array}{ccccccccc}
 & & 1 & & 1 & & 1 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
1 & \longrightarrow & H \cap Z & \longrightarrow & Z & \longrightarrow & Z/H \cap Z & \longrightarrow & 1 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
1 & \longrightarrow & H & \longrightarrow & \mathrm{SL}_n(\mathbf{F}) & \longrightarrow & \mathrm{SL}_n(\mathbf{F})/H & \longrightarrow & 1 \\
 & & \downarrow & & {\scriptstyle\eta}\downarrow & & \downarrow & & \\
1 & \longrightarrow & \eta(H) & \longrightarrow & \mathrm{PSL}_n(\mathbf{F}) & \longrightarrow & \mathrm{PSL}_n(\mathbf{F})/\eta(H) & \longrightarrow & 1 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 1 & & 1 & & 1 & &
\end{array}$$

has exact rows and columns. Now, either $\eta(H) = 1$ or $\eta(H) = \mathrm{PSL}_n(\mathbf{F})$, since $\mathrm{PSL}_n(\mathbf{F})$ is simple. If $\eta(H) = 1$ then $H \cap Z = H$, so that $H \subset Z$. If $\eta(H) = \mathrm{PSL}_n(\mathbf{F})$ then $Z/H \cap Z \cong \mathrm{SL}_n(\mathbf{F})/H$ is abelian, hence trivial, since $\mathrm{SL}_n(\mathbf{F})$ is perfect (i.e. $\mathrm{SL}_n(\mathbf{F})_{ab}$ is trivial), so that $H = \mathrm{SL}_n(\mathbf{F})$. □

Observe that the exact sequence of groups

$$1 \to \mathrm{SL}_n(\mathbf{F}) \to \mathrm{GL}_n(\mathbf{F}) \xrightarrow{\det} \mathbf{F}^* \to 1$$

splits; for example the homomorphism $s : \mathbf{F}^* \to \mathrm{GL}_n(\mathbf{F})$ defined by $s(x) = x \oplus I_{n-1}$ is a section. Thus $\mathrm{GL}_n(\mathbf{F}) \cong \mathrm{SL}_n(\mathbf{F}) \bowtie \mathbf{F}^*$, the semidirect product, where the action $\alpha : \mathbf{F}^* \times \mathrm{SL}_n(\mathbf{F}) \to \mathrm{SL}_n(\mathbf{F})$ is given by $\alpha(x, A) = s(x)As(x)^{-1}$.

Proposition 1.2. *Let G be a normal subgroup of $\mathrm{GL}_n(\mathbf{F})$, where $n \neq 2$ and $|\mathbf{F}| \neq 2, 3$. Then, either*

(1) *G consists of scalar matrices and therefore $G \subset Z(\mathrm{GL}_n(\mathbf{F})) \cong \mathbf{F}^*$, or*
(2) *G contains a non-scalar matrix and is a semidirect product $G \cong \mathrm{SL}_n(\mathbf{F}) \bowtie U$, where $U = \det(G) \subset \mathbf{F}^*$.*

Proof. The commutative diagram

$$\begin{array}{ccccccccc}
 & & 1 & & 1 & & 1 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
1 & \longrightarrow & G \cap \mathrm{SL}_n(\mathbf{F}) & \longrightarrow & G & \xrightarrow{\det} & U & \longrightarrow & 1 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
1 & \longrightarrow & \mathrm{SL}_n(\mathbf{F}) & \longrightarrow & \mathrm{GL}_n(\mathbf{F}) & \xrightarrow{\det} & \mathbf{F}^* & \longrightarrow & 1
\end{array}$$

has exact rows and columns. The bottom sequence is split by the homomorphism $s : \mathbf{F}^* \to \mathrm{GL}_n(\mathbf{F})$ defined by $s(x) = x \oplus I_{n-1}$, so that $\mathrm{GL}_n(\mathbf{F}) \cong \mathrm{SL}_n(\mathbf{F}) \rtimes \mathbf{F}^*$. The action $\alpha : \mathbf{F}^* \times \mathrm{SL}_n(\mathbf{F}) \to \mathrm{SL}_n(\mathbf{F})$ is given by $\alpha(x, A) = s(x)As(x)^{-1}$.

If G consists of scalar matrices then the assertion is obvious. If G contains a non-scalar matrix A then for some $S \in \mathrm{SL}_n(\mathbf{F})$ the element $[S, A] = SAS^{-1}A^{-1}$ of $G \cap \mathrm{SL}_n(\mathbf{F})$ is not scalar. For, suppose to the contrary that $[S, A] = SAS^{-1}A^{-1} = \lambda_S I$, i.e. $SAS^{-1} = \lambda_S A$, for all $S \in \mathrm{SL}_n(\mathbf{F})$. Then $\lambda : \mathrm{SL}_n(\mathbf{F}) \to \mathbf{F}^*$ is a homomorphism of groups and in particuar $\lambda_{[S,T]} = 1$ for all $S, T \in \mathrm{SL}_n(\mathbf{F})$. Since $\mathrm{SL}_n(\mathbf{F})$ is perfect, i.e. $[\mathrm{SL}_n(\mathbf{F}), \mathrm{SL}_n(\mathbf{F})] = \mathrm{SL}_n(\mathbf{F})$, it follows that $\lambda_S = 1$ and hence $[S, A] = I$ for all $S \in \mathrm{SL}_n(\mathbf{F})$, which means that A is scalar. Thus, if G contains a non-scalar matrix then so does $G \cap \mathrm{SL}_n(\mathbf{F})$, and $G \cap \mathrm{SL}_n(\mathbf{F}) = \mathrm{SL}_n(\mathbf{F})$ by Lemma 1.1. Then $\det^{-1}(U) = G$, hence the top exact sequence of the diagram splits, and $G \cong \mathrm{SL}_n(\mathbf{F}) \rtimes U$. $\square$

Corollary 1.3. *If $n \neq 2$ anf $|\mathbf{F}| \neq 2, 3$ then the subgroup G of $\mathrm{GL}_n(\mathbf{F})$ generated by the similarity orbit of a non-scalar invertible matrix A is of the form $G \cong \mathrm{SL}_n(\mathbf{F}) \rtimes U$, where U is the cyclic subgroup of $\mathbf{F}^*$ generated by $\det A$.*

To determine the semigroup $\mathcal{S}$ (as opposed to the group) generated by the similarity orbit of an invertible matrix is more complicated. Since every square matrix has a rational canonical form it is useful to start with the companion matrix of a polynomial, i.e. a cyclic matrix.

Lemma 1.4. *The semigroup $\mathcal{S}$ in $\mathrm{GL}_n(\mathbf{F})$ generated by the similarity orbit of the companion matrix A of the polynomial $p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ with $\det A = a_0 \neq 0$ contains the diagonal matrix $I_{n-1} \oplus a_0^2$ and the scalar matrix $a_0^2 I$.*

Proof. If Q is the involution obtained from the identity I by reversing the order of the rows then $B = QAQ$ is the matrix obtained from A by first reversing the order of the rows of A to get a matrix C and then reversing the order of the columns of C to get B. Then

$$BA = \begin{pmatrix} I_{n-1} & X \\ 0 & a_0^2 \end{pmatrix}$$

for some X, where I_{n-1} is the identity matrix of size $n-1$. If $a_0^2 \neq 1$ then BA is similar to $I_{n-1} \oplus a_0^2$, as can be seen by replacing the last vector in the standard ordered basis $\{e_i | 1 \le i \le n\}$ of $\mathbf{F}^n$ by $e_n + (1/(a_0^2 - 1)) \sum_{i=1}^{n-1} x_i e_i$. Thus we are done in this case, since by a cyclic permuation similarity argument $a_0^2 I$ is in $\mathcal{S}$. If $a_0^2 = 1$ then $BA = I + N$ with $N^2 = 0$. Since $I + N$ is similar to $I - N$, which is easily seen by replacinge$_n$ by $-e_n$ in the standard ordered basis of $\mathbf{F}^n$, it follows that $(I + N)(I - N) = I$ is in $\mathcal{S}$. $\square$

Proposition 1.5. *The semigroup $\mathcal{S}$ in $\mathrm{GL}_n(\mathbf{F})$ generated by the similarity orbit of an invertible matrix A contains an upper-triangular matrix U with $\det U = \det A^2$, a diagonal matrix D with $\det D = \det A^4$ and a non-zero scalar matrix λI with $\lambda = \det A^{4n}$.*

Proof. We may assume without loss of generality that A is in rational canonical form. Apply Lemma 1.4 to each companion matrix in the rational decomposition of A to get an upper-triangular matrix $BA \simeq (I+N) \oplus D$, where B is similar to A, D is diagonal, $\det D = \det A^2$ and $N^2 = 0$. Again, since $(I + N) \oplus D$ is similar to $(I - N) \oplus D$, it follows that $(I + N) \oplus D)((I - N) \oplus D) = I \oplus D^2$. Cyclicly permuting the diagonal entries of $I \oplus D^2$ yields n mutually similar diagonal matrices. The product of these diagonal matrices is the scalar matrix $\lambda I \in \mathcal{S}$, where $\lambda = \det A^{4n}$. $\square$

Corollary 1.6. *Let $\mathcal{S}$ be the semigroup in $\mathrm{GL}_n(\mathbf{F})$ generated by the similarity orbit of an invertible matrix A. Then $\mathcal{S}$ is a normal subgroup of $\mathrm{GL}_n(\mathbf{F})$ if and only if $\det A$ is a root of unity. If $d = \det A$ is a root of unity and A is not scalar then $\mathcal{S} \cong \mathrm{SL}_n(\mathbf{F}) \bowtie < d >$, except when $n = 2$ and $|\mathbf{F}| = 2, 3$. In particular, if $d = 1$ then $\mathcal{S} = \mathrm{SL}_n(\mathbf{F})$, except when $n = 2$ and $|\mathbf{F}| = 2, 3$*

Proof. If $d = \det A$ is not a root of unity then $\det S \neq 1$ for all $S \in \mathcal{S}$ and the semigroup $\mathcal{S}$ is not a subgroup of $\mathrm{GL}_n(\mathbf{F})$. If $d^m = 1$ then $I = D^m = XSAS^{-1}$ in $\mathrm{SL}_n(\mathbf{F})$ for some $X \in \mathrm{SL}_n(\mathbf{F})$ and some $S \in \mathrm{GL}_n(\mathbf{F})$, where D is the diagonal matrix of Proposition 1.5. Thus, $A^{-1} = S^{-1}XS \in \mathcal{S}$ and $\mathcal{S}$ is a subgroup of $\mathrm{GL}_n(\mathbf{F})$. Now apply Proposition 1.2. $\square$

In the two exceptional cases $n = 2$ and $|\mathbf{F}| = 2, 3$ the group $\mathrm{PSL}_n(\mathbf{F})$ is not simple and $\mathrm{SL}_n(\mathbf{F})$ is not perfect. These cases have to be considered separately.

The group $\mathrm{GL}_2(\mathbf{Z}_2)$ is not abelian and $|\,\mathrm{GL}_2(\mathbf{Z}_2)| = 6$, so that $\mathrm{PSL}_2(\mathbf{Z}_2) \cong \mathrm{SL}_2(\mathbf{Z}_2) \cong \mathrm{GL}_2(\mathbf{Z}_2) \cong S_3$, the symmetric group on three symbols. The only proper normal subgroup of $\mathrm{GL}_2(\mathbf{Z}_2)$ is therefore the cyclic subgroup C_3 of order 3 generated by

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{or its inverse} \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Proposition 1.7. *If $I \neq A \in \mathrm{GL}_2(\mathbf{Z}_2)$ then $\mathcal{S} \cong C_3$ if A has order 3 and $\mathcal{S} = \mathrm{GL}_2(\mathbf{Z}_2)$ otherwise.* $\square$

In the case of $\mathrm{GL}_2(\mathbf{Z}_3)$ we have $|\,\mathrm{GL}_2(\mathbf{Z}_3)| = 48$ and $Z(\mathrm{GL}_2(\mathbf{Z}_3)) \cong C_2$ is the cyclic subgroup of order 2 generated by $2I$. In the commutative diagram with exact rows and columns

$$\begin{array}{ccccc} Z & = & Z & & \\ \downarrow & & \downarrow & & \\ \mathrm{SL}_2(\mathbf{Z}_3) & \longrightarrow & \mathrm{GL}_2(\mathbf{Z}_3) & \xrightarrow{\det} & \mathbf{Z}_3^* \\ \eta\downarrow & & \eta\downarrow & & \| \\ \mathrm{PSL}_2(\mathbf{Z}_3) & \longrightarrow & \mathrm{PGL}_2(\mathbf{Z}_3) & \xrightarrow{\det} & \mathbf{Z}_3^* \end{array}$$

the determinant map is split by the homomorphism $s : \mathbf{Z}_3^* \to \mathrm{GL}_2(\mathbf{Z}_3)$ defined by $s(2) = \mathrm{diag}[2, 1]$. Moreover, the Sylow 2-subgroups of $\mathrm{SL}_2(\mathbf{Z}_3)$ and $\mathrm{PSL}_2(\mathbf{Z}_3)$ are normal, they are a copy of the quaternion group Q generated by the two matrices

$$X = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

and a copy of the Klein 4-group V generated by $\eta(X)$ and $\eta(Y)$, respectively. We have a commutative diagram with exact rows and columns

$$\begin{array}{ccccc} Z & = & Z & & \\ \downarrow & & \downarrow & & \\ Q & \longrightarrow & \mathrm{SL}_2(\mathbf{Z}_3) & \xrightarrow{p} & C_3 \\ \eta\downarrow & & \eta\downarrow & & \| \\ V & \longrightarrow & \mathrm{PSL}_2(\mathbf{Z}_3) & \longrightarrow & C_3 \end{array}$$

in which the canonical projection $p : \mathrm{SL}_2(\mathbf{Z}_3) \to C_3$ is split by the homomorphism $t : C_3 \to \mathrm{SL}_2(\mathbf{Z}_3)$, where

$$t(x) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

the image of a generator of C_3, generates a Sylow 3-subgroup of order 3 in $\mathrm{SL}_2(\mathbf{Z}_3)$. Observe that $s(\mathbf{Z}_3^*)$ acts on $t(C_3)$ and on Q while $t(C_3)$ acts on Q by conjugation, so that $C_3 \rtimes \mathbf{Z}_3^* \cong S_3$. Thus, $\mathrm{PSL}_2(\mathbf{Z}_3) \cong V \rtimes C_3$, $\mathrm{SL}_2(\mathbf{Z}_3) \cong Q \rtimes C_3$ and $\mathrm{GL}_2(\mathbf{Z}_3) \cong \mathrm{SL}_2(\mathbf{Z}_3) \rtimes \mathbf{Z}_3^* \cong Q \rtimes S_3$. There are three Sylow 2-subgroups of order 16 in $\mathrm{GL}_2(\mathbf{Z}_3)$, namely $Q \rtimes \mathbf{Z}_3^*$ and its conjugates. They intersect in the normal subgroup Q. The proper normal subgroups of $\mathrm{GL}_2(\mathbf{Z}_3)$ are therefore $Z \cong C_2$, Q and $\mathrm{SL}_2(\mathbf{Z}_3)$.

Proposition 1.8. *Let* $I \neq A \in \mathrm{GL}_2(\mathbf{Z}_3)$.

(1) *If* $\det A = 1$ *then* $\mathcal{S} = Z, Q, \mathrm{SL}_2(\mathbf{Z}_3)$ *depending on whether the order of* A *is* 2, 4 *or divisible by* 3.

(2) *If* $\det A = 2$ *then* $\mathcal{S} = \mathrm{GL}_2(\mathbf{Z}_3)$. □

The main result of [S] will be used repeatedly in the next section. We record it here, without proof, for future reference.

Theorem 1.9. *Let $A \in \mathrm{GL}_n(\mathbf{F})$ be nonscalar and let β_j, γ_j $(1 \le j \le n)$ be elements of $\mathbf{F}^*$ such that $\prod_{j=1}^n \beta_j\gamma_j = \det A$. Then there exist matrices B and C in $\mathrm{GL}_n(\mathbf{F})$ with eigenvalues $\beta_1, \dots, \beta_n$ and $\gamma_1, \dots, \gamma_n$, respectively, such that $A = BC$. Furthermore, B and C can be chosen so that B is lower triangularizable and C is simultaneously upper triangularizable.* □

2. Some bounds on the number of similarity factors required

In this section we have to assume that the field $\mathbf{F}$ has enough elements, $|\mathbf{F}| > 2n$ should suffice. The following result of Cater [C], which we quote here without proof, will be used in our considerations.

Lemma 2.1. *If M is a non-scalar in $\mathrm{GL}_n(\mathbf{F})$ and $\det M = x_1x_2\dots x_n$ then there is a factorization $M = A_1A_2\dots A_n$ with $\det A_i = x_i$ and $\mathrm{rank}(A_i - I) = 1$ for $i = 1, 2, \dots, n$.* □

Observe that the properties of the matrices A_i of Lemma 2.1 imply that A_i is similar to $(I_2 + J_2) \oplus I_{n-2}$ if $x_i = 1$ and similar to $x_i \oplus I_{n-1}$ if $x_i \ne 1$. Here is an immediate consequence of Cater's result.

Proposition 2.2. *Let A be a non-scalar element of $\mathrm{GL}_n(\mathbf{F})$ such that $\mathrm{rank}(A - I) = 1$. If $\det T = \det A^n$ then $T = A_1A_2\dots A_n$, where A_i is similar to A for $i = 1, 2, \dots, n$.*

Proof. The conditions imposed on A imply that A is similar to $(I_2 + J_2) \oplus I_{n-2}$ if $\det A = 1$ and similar to $\det A \oplus I_{n-1}$ if $\det A \ne 1$. By Cater's Lemma 2.1 we see that $T = A_1A_2\dots A_n$, where $\det A_i = \det A$ and $\mathrm{rank}(A_i - I) = 1$, and hence where A_i is similar to A for $i = 1, 2, \dots, n$. □

Corollary 2.3. *If T is in $\mathrm{SL}_n(\mathbf{F})$ then $T = A_1A_2\dots A_k$ for some k such that $0 \le k \le n$, where A_i is similar to $A = (I_2 + J_2) \oplus I_{n-2}$ for $i = 1, 2, \dots, k$.* □

Lemma 2.4. *If $A \in GL_n(\mathbf{F})$ is cyclic, then every $T \in GL_n(\mathbf{F})$ with distinct eigenvalues and $\det T = \det A^2$ has a factorization $T = A_1A_2$, where A_i is similar to A for $i = 1, 2$.*

Proof. The matrix A is similar to the companion matrix of its characteristic polynomial $p(x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1} + x^n$. Thus we may assume that

$$A = \begin{pmatrix} 0 & & & a_0 \\ 1 & & & a_1 \\ & \ddots & & \vdots \\ & & 1 & a_{n-1} \end{pmatrix}.$$

It is easy to see that via a suitable diagonal similarity A is similar to a matrix of the form

$$B = \begin{pmatrix} 0 & & & b_0 \\ x_1 & & & b_1 \\ & \ddots & & \vdots \\ & & x_{n-1} & a_{n-1} \end{pmatrix},$$

where $x_1, x_2, \dots, x_{n-1}$ can be chosen arbitrarily in $\mathbf{F}^*$, and where the determinant condition $b_0 x_1 x_2 \dots x_{n-1} = a_0$ holds. Then

$$S = \begin{pmatrix} a_{n-1} & x_{n-1} & & \\ \vdots & & \ddots & \\ b_1 & & & x_1 \\ b_0 & & & 0 \end{pmatrix} \begin{pmatrix} 0 & & & a_0 \\ 1 & & & a_1 \\ & \ddots & & \vdots \\ & & 1 & a_{n-1} \end{pmatrix} = \begin{pmatrix} x_{n-1} & & & * \\ & \ddots & & \vdots \\ & & x_1 & \\ & & & b_0 a_0 \end{pmatrix}$$

is upper-triangular, and the first factor of S is similar to B via the similarity given by the involution obtained by reversing the order of the rows of the identity matrix. Since $x_1, x_2, \dots, x_{n-1}$ and $b_0 a_0$ can be taken to be the distinct eigenvalues of T we conclude that T is similar to S, and thus T is of the desired form. $\square$

Proposition 2.5. *Suppose that $|F| > 2n$. If the rational canonical form of A has no scalar direct summand then there exists a $T \in \mathrm{GL}_n(\mathbf{F})$ with distinct eigenvalues and* $\det T = \det A^2$ *such that $T = A_1 A_2$ with A_1 and A_2 similar to A. Furthermore, the eigenvalues of T can be chosen outside a given subset E of $\mathbf{F}^*$ if $|F| \geq 2(|E| + n)$.*

Proof. Assume without loss of generality that A is in rational canonical form $A = R_1 \oplus R_2 \oplus \dots \oplus R_m$. By hypothesis each rational cell R_i has size $k_i \geq 2$. Let $n_0 = 0$ and $n_j = n_{j-1} + k_j$ for $j = 1, 2, \dots, m$. We want to apply Lemma 2.4 in sequence to each rational cell R_i. First choose $n_1 - 2$ distinct elements $x_1, \dots, x_{n_1-2}$ of $\mathbf{F}^*$ outside E. Then choose distinct elements x_{n_1-1} and x_{n_1} outside $E' = E \cup \{x_1, \dots, x_{n_1-2}\}$ such that $x_1 x_2 \dots x_{n_1} = \det R_1^2$. This is possible if $|F^*| > 2|E'| + 2 = 2(|E| + n_1 - 1)$. We have now used n_1 distinct elements of F^*. Now let $E_1 = E \cup \{x_1, \dots, x_{n_1}\}$, and choose in the same way distict elements $x_{n_1+1}, \dots, x_{n_2}$ of F^* outside E_1 such that $x_{n_1+1} \dots x_{n_2} = \det R_2^2$. This is possible if $|F^*| > 2(|E_1| + k_2 - 1) = 2(E| + n_2 - 1)$. Continue this process to obtain a sequence $\{x_1, x_2, \dots, x_n\}$ of distinct elements of $\mathbf{F}^*$ outside E with $x_{n_j+1} \dots x_{n_{j+1}} = \det R_j^2$ for $j = 0, 1, \dots, m-1$. This is possible if $|F^*| > 2(|E| + n - 1)$. Now let $T_j = \mathrm{diag}[x_{n_j+1}, \dots, x_{n_{j+1}}]$. Applying Lemma 2.4, we get factorizations $T_j = R_j' R_j''$ with R_j' and R_j'' each similar to R_j. Then $T = T_1 \oplus T_2 \oplus \dots \oplus T_m = R' R''$, where $R' = R_1' \oplus R_2' \oplus \dots \oplus R_m'$ and $R'' = R_1'' \oplus R_2'' \oplus \dots \oplus R_m''$ are both similar to A. $\square$

Theorem 2.6. *If the rational canonical form of A has no scalar direct summand then every matrix B with* $\det B = \det A^4$ *is of the form $B = A_1 A_2 A_3 A_4$, where A_i is similar to A for $i = 1, 2, 3, 4$.*

Proof. Use Theorem 1.9 to write $B = LU$, where L is lower-triangular and U is upper-triangular, each with the same spectrum as the operator T of Proposition 2.5.

Thus L and U are both similar to T. It then follows from Proposition 2.5 that $B = LU = A_1A_2A_3A_4$, where A_i is similar to A for $i = 1, 2, 3, 4$. □

Corollary 2.7. *Let $A \in \mathrm{GL}_n(\mathbf{F})$ be such that its rational canonical form has no scalar direct summand, and let k be any natural number. Then every matrix $B \in \mathrm{GL}_n(\mathbf{F})$ with $\det B = \det A^{4k}$ is of the form $B = A_1A_2 \dots A_{4k}$, where A_i is similar to A for $i = 1, 2, \dots, 4k$.* □

Corollary 2.8. *If the rational canonical form of $A \in \mathrm{SL}_n(\mathbf{F})$ has no scalar direct summand then every matrix $B \in \mathrm{SL}_n(\mathbf{F})$ is of the form $B = A_1A_2A_3A_4$ where A_i is similar to A for $i = 1, 2, 3, 4$.*

For a matrix $A \in \mathrm{GL}_n(\mathbf{F})$ whose rational canonical form has a scalar direct summand of size one the bound on the similarity factors depends on the multiplicity of this summand. The 'worst' case occurs when that scalar direct summand has multiplicity $n-2$, i.e. when A is diagonalizable with an eigenvalue of multiplicity $n-1$.

Theorem 2.9. *If the rational canonical form of $A \in \mathrm{GL}_n(\mathbf{F})$ has a scalar direct summand of multiplicity $r - 1 \leq n - 2$ then every non-scalar $T \in \mathrm{GL}_n(\mathbf{F})$ with $\det T = \det A^{4r}$ is of the form $T = A_1A_2 \dots A_{4r}$, where A_i is similar to A for $i = 1, 2, \dots, 4r$.*

Proof. Without loss of generality we may assume that the matrix A is in rational canonical form $A = cI_{r-1} \oplus R_1 \oplus \dots \oplus R_m$, where each rational cell R_j has size at least 2. Apply Proposition 2.5 with $E = \{c^2\}$ to $R_1 \oplus R_2 \oplus \dots \oplus R_m$ to get a matrix

$$B = A_1A_2 = c^2 I_{r-1} \oplus \mathrm{diag}[d_0, d_1, \dots, d_{n-r}] = D_0 \oplus \mathrm{diag}[d_1, d_2, \dots, d_{n-r}] = D_0 \oplus D_1$$

so that the entries $c^2, d_0, d_1, \dots, d_{n-r}$ are all distinct, with A_1 and A_2 similar to A. This is possible if $|F^*| > 2(n-r)$. Then $D_0 = c^2 I_{r-1} \oplus d_0$ and $\mathrm{rank}(\frac{1}{c^2}D_0 - I_r) = 1$. Setting $\alpha = (-1)^{r-1}(d_0/c^2)^r$ and applying Lemma 2.1 we conclude that

$$\begin{pmatrix} & & & \alpha \\ 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \end{pmatrix} = M_1M_2 \dots M_r$$

with $\det M_i = d_0/c^2 \neq 1$ and $\mathrm{rank}(M_i - I_r) = 1$. Thus M_i is similar to $I_{r-1} \oplus \frac{d_0}{c^2} = \frac{1}{c^2}D_0$. Multiplying by c^{2r} we get the matrix

$$P = \begin{pmatrix} & & & c^{2r}\alpha \\ c^{2r} & & & 0 \\ & \ddots & & \vdots \\ & & c^{2r} & 0 \end{pmatrix} = P_1P_2 \dots P_r$$

with $P_i = c^2 M_i$ similar to D_0 for $i = 1, 2, \ldots, r$. Moreover, by repeated applications of Theorem 1.7 we can find a diagonal matrix $Q = \mathrm{diag}[q_1, q_2, \ldots, q_{n-r}]$ with distinct diagonal entries, distinct from the eigenvalues of P, such that $\det Q = \det D_1^r$ and $Q = Q_1 Q_2 \ldots Q_r$, where Q_i is similar to D_1 for $i = 1, 2, \ldots, r$. Thus, $C = P \oplus Q$ is cyclic, $\det C = \det(P)\det(Q) = \det B^r$ and $C = B_1 B_2 \ldots B_r = A_1 A_2 \ldots A_{2r}$, where $B_i = P_i \oplus Q_i$ is similar to $B = D_0 \oplus D_1$ for $i = 1, 2, \ldots, r$ and A_j is similar to A for $j = 1, 2, \ldots, 2r$.

Thus, by Theorem 1.9, every matrix $T \in \mathrm{GL}_n(\mathbf{F})$ with $\det T = \det C^2 = \det B^{2r} = \det A^{4r}$ is of the form

$$T = C_1 C_2 = B_1 B_2 \ldots B_{2r} = A_1 A_2 \ldots A_{4r}$$

with C_i is similar to C, B_j is similar to B and A_k is similar to A. □

Corollary 2.10. *If $A \in \mathrm{GL}_n(\mathbf{F})$ is not scalar and $s = \mathrm{lcm}(1, 2, \ldots, n-1)$, then every $T \in \mathrm{GL}_n(\mathbf{F})$ with $\det T = \det A^{4s}$ is of the form $T = A_1 A_2 \ldots A_{4s}$.* □

3. Semigroups generated by the similarity orbit of a singular matrix

We first prove a preliminary result for the similarity semigroup when $\operatorname{rank} A = n-1$ and then apply it to to show that in the general when $\operatorname{rank} A < n$ the similarity semigroup of A consists of all matrices of rank less than or equal to $\operatorname{rank} A$.

Proposition 3.1. *The semigroup in $M_n(F)$ generated by the similarity orbit of a matrix A with $\operatorname{rank} A = n-1$ consists of all matrices of rank less than or equal to $n-1$.*

Proof. Let $\mathcal{S}$ be the semigroup generated by the similarity orbit of the matrix A of rank $n-1$ in $M_n(\mathbf{F})$. The proof will be in four steps.

Step 1) We first show that $\mathcal{S}$ contains a matrix $C = X \oplus 0$ for some invertible X of size $n-1$. By Fitting's Lemma, see for example [B], we have $\mathbf{F}^n = \operatorname{im} A^m \oplus \ker A^m$ for some natural number m, so that we may assume that $A = Y \oplus N$, where Y is invertible and N is nilpotent in Jordan canonical form. Then $B = Y \oplus N^T$ is similar to A and $AB = Y^2 \oplus I \oplus 0 = X \oplus 0$, where X is invertible of size $n-1$.

Step 2) Next we can prove that $\mathcal{S}$ contains a matrix $Y = \lambda I_{n-1} \oplus 0$, where $\lambda \neq 0$ and I_{n-1} is the identity matrix of rank $n-1$. In the matrix $C = X \oplus 0$ of step 1) the matrix X is invertible and we can get the result by applying Proposition 1.5.

Step 3) Now we show that $\mathcal{S}$ contains for each $r = 0, 1 \ldots, n-1$ a matrix of the form $\lambda I_r \oplus N$, where N is nilpotent of maximal rank $n-r-1$. This is certainly true for $r = n-1$ by step 2). If $r = n-2$ and $Y = \lambda I_{n-1} \oplus 0$ is the matrix obtained in step 2) then

$$\begin{pmatrix} \lambda & & & \\ & \ddots & & \\ & & \lambda & \\ & & -\lambda & \end{pmatrix} \begin{pmatrix} \lambda & & & \\ & \ddots & & \\ & & \lambda & \lambda \\ & & & \end{pmatrix} = \begin{pmatrix} \lambda^2 & & & \\ & \ddots & & \\ & & \lambda^2 & \lambda^2 \\ & & -\lambda^2 & -\lambda^2 \end{pmatrix},$$

that is

$$Q^{-1}YQS^{-1}YS = \lambda^2 I_{n-2} \oplus \lambda^2 \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix},$$

which is similar to

$$\begin{pmatrix} \lambda^2 I_{n-2} & & \\ & 0 & \lambda^2 \\ & 0 & 0 \end{pmatrix} = \lambda^2 I_{n-2} \oplus \lambda^2 J_2.$$

Here we used the similarities

$$Q^{-1}YQ = Q^{-1}Y = \lambda I_{n-2} \oplus \begin{pmatrix} \lambda & 0 \\ -\lambda & 0 \end{pmatrix} \quad \text{and} \quad S^{-1}YS = YS = \lambda I_{n-2} \oplus \begin{pmatrix} \lambda & \lambda \\ 0 & 0 \end{pmatrix},$$

where the elementary matrix $Q = E_{n,n-1}$ is obtained from I_n by adding the $(n-1)$-th row to the n-th row and $S = Q^T$ is the transpose.

Now proceed by backward induction on r using

$$\lambda^{2(n-r-1)} \begin{pmatrix} I_r & \\ & J_{n-r} \end{pmatrix} \lambda^2 \begin{pmatrix} I_{r-1} & & \\ & J_2 & \\ & & I_{n-r-1} \end{pmatrix} = \lambda^{2(n-r)} \begin{pmatrix} I_{r-1} & 0 \\ 0 & J_{n-r+1} \end{pmatrix}$$

which is the same as

$$\lambda^{2(n-r-1)}(I_r \oplus J_{n-r})\lambda^2(I_{r-1} \oplus J_2 \oplus I_{n-r-1}) \simeq \lambda^{2(n-r)}(I_{r-1} \oplus J_{n-r+1}),$$

or the same as

$$(I_r \oplus J_{n-r})(I_{r-1} \oplus J_2 \oplus I_{n-r-1}) = I_{r-1} \oplus J_{n-r+1},$$

where J_s is the nilpotent Jordan cell of size s and rank $s-1$.

Sep 4) Finally we prove that $\mathcal{S}$ contains every matrix of the form $Z \oplus 0$ for every invertible matrix Z of size $n-1$. By step 3) the big Jordan cell J_n is in $\mathcal{S}$ and so are its transpose J_n^t and all their powers. Moreover $J_n J_n^t = I_{n-1} \oplus 0$ is idempotent of rank $n-1$ and $J_n^k (J_n^T)^k = I_{n-k} \oplus O_k$ is idempotent of rank $n-k$. Thus $\mathcal{S}$ contains all idempotents of rank less than or equal to $n-1$. Then

$$\begin{pmatrix} I_{n-1} & x \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_{n-1} & 0 \\ y^t & 0 \end{pmatrix} = \begin{pmatrix} I_{n-1} + xy^t & 0 \\ 0 & 0 \end{pmatrix}$$

yields the result. This is all we need to proceed with the general case when rank $A < n$ and the final argument is done in the proof of the next theorem. □

Theorem 3.2. *The semigroup $\mathcal{S}$ in $M_n(\mathbf{F})$ generated by the similarity orbit of a matrix A of rank $r < n$ consists of all matrices of rank $\leq r$.*

Proof. Let $\operatorname{rank} A = r = n - u$. The argument used in step 1) of Proposition 3.1 shows that $\mathcal{S}$ contains a matrix of the form $X \oplus O_u$ for some invertible matrix X of size r. That $\mathcal{S}$ contains a matrix $Y = \lambda I_r \oplus O_u$ for some scalar $\lambda \neq 0$ again follows from Proposition 1.5 as in step 2) of Proposition 3.1. As in step 3) of Proposition 3.1 with $n = r + 1$ we show that for each $s = 0, 1, \ldots, r$ the semigroup $\mathcal{S}$ contains a matrix of the form $\lambda I_s \oplus N \oplus O_{u-1}$, where $N \simeq J_{r-s+1}$ is nilpotent of maximal rank $r - s$. As in step 4) of Proposition 3.1 it now follows that $\mathcal{S}$ contains all matrices of the form $Z \oplus O_u$ for every invertible matrix Z of size r.

This shows in particular that $K = J_{r+1} \oplus O_{u-1}$, all its powers and their transposes are in $\mathcal{S}$. But then $K^l(K^l)^T = I_{r-l} \oplus O_{n-r-l}$ is in $\mathcal{S}$ for $l = 1, 2, \ldots r$, and hence $\mathcal{S}$ contains all idempotents of rank $\leq r$, and hence all matrices of the form $C \oplus O_w$ for invertible C and $u \leq w \leq n$.

Now we want to prove that if $B \in M_n(\mathbf{F})$ and $\operatorname{rank}(B) = v \leq r$ then $B \in \mathcal{S}$. By Fitting's Lemma $B \simeq B_0 \oplus N$, where B_0 is invertible of size $s \geq 0$ and N is nilpotent of rank $v - s$. More precisely,

$$B \simeq B_0 \oplus N \simeq B_0 \oplus J_{s_1} \oplus J_{s_2} \oplus \ldots \oplus J_{s_t} \oplus O_w =$$
$$(B_0 \oplus (I_{s_1-1} \oplus 0) \oplus \ldots \oplus (I_{s_t-1} \oplus 0) \oplus O_w)(I_s \oplus J_{s_1} \oplus \ldots \oplus J_{s_t} \oplus O_w)$$

when N is in Jordan form. Since $n = s + s_1 + s_2 + \ldots + s_t + w = v + t + w$ it follows that the number of Jordan cells is $t = n - v - w \leq n - v$. The first factor on the right is similar to $B_0 \oplus I_{v-s} \oplus O_{w+t}$, hence belongs to $\mathcal{S}$. The second factor is in the semigroup generated by the similarity orbit of

$$I_v \oplus O_{w+t} \simeq I_s \oplus (I_{s_1-1} \oplus 0) \oplus (I_{s_2-1} \oplus 0) \oplus \ldots \oplus (I_{s_t-1} \oplus 0) \oplus O_w \in \mathcal{S},$$

since J_{s_j} is in the semigroup generated by the similarity orbit of $I_{s_j-1} \oplus 0$ in $M_{s_j}(\mathbf{F})$ for $j = 1, 2, \ldots, t$ by step 3) in the proof of Proposition 3.1. This proves that B is in $\mathcal{S}$. $\square$

References

[AB] J.L. Alperin, R.B. Bell. *Groups and Representations.* Graduate Texts in Mathematics 162, Springer Verlag, 1995

[B] D. Benson. *Representations and cohomology I.* Cambridge studies in advanced mathematics 30, 1991

[C] F.S. Cater. Products of Central Collineations. *Lin. Alg. and Appl.* 19, 251-274, 1978.

[S] A. Sourour. A Factorization Theorem for Matrices. *Lin. Multilin. Alg.* 19, 141-147, 1986.

L. Grunenfelder and H. Radjavi : Department of Mathematics and Statistics, Dalhousie University, Halifax, Nova Scotia, Canada, B3H 3J5

M. Omladič : Department of Mathematics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia

A. Sourour : Department of Mathematics and Statistics, University of Victoria, Victoria, B.C. Canada