Canad. Math. Bull. Vol. **37** (3), 1994 pp. 374–383

LINEAR OPERATORS PRESERVING SIMILARITY CLASSES AND RELATED RESULTS

CHI-KWONG LI AND STEPHEN PIERCE

ABSTRACT. Let M_n be the algebra of $n \times n$ matrices over an algebraically closed field $\mathbb{F}$ of characteristic zero. For $A \in M_n$, denote by $\mathcal{S}(A)$ the collection of all matrices in M_n that are similar to A. In this paper we characterize those invertible linear operators ϕ on M_n that satisfy $\phi(\mathcal{S}) \subseteq \mathcal{S}$ or $\phi(\bar{\mathcal{S}}) \subseteq \bar{\mathcal{S}}$, where $\mathcal{S} = \mathcal{S}(A_1) \cup \cdots \cup \mathcal{S}(A_k)$ for some given $A_1, \ldots, A_k \in M_n$ and $\bar{\mathcal{S}}$ denotes the (Zariski) closure of $\mathcal{S}$. Our theorem covers a result of Howard on linear operators mapping the set of matrices annihilated by a given polynomial into itself, and extends a result of Chan and Lim on linear operators commuting with the function $f(x) = x^k$ for a given positive integer $k \geq 2$. The possibility of weakening the invertibility assumption in our theorem is considered, a partial answer to a conjecture of Howard is given, and some extensions of our result to arbitrary fields are discussed.

1. **Introduction.** Let M_n be the algebra of $n \times n$ matrices over an algebraically closed field $\mathbb{F}$ of characteristic zero. For $A \in M_n$, denote by $\mathcal{S}(A)$ the *similarity class* of A, *i.e.*, the collection of all matrices in M_n that are similar to A. In this paper we characterize those invertible linear operators ϕ on M_n that satisfy $\phi(\mathcal{S}) \subseteq \mathcal{S}$ or $\phi(\bar{\mathcal{S}}) \subseteq \bar{\mathcal{S}}$, where $\mathcal{S} = \mathcal{S}(A_1) \cup \cdots \cup \mathcal{S}(A_k)$ for some given $A_1, \ldots, A_k \in M_n$ and $\bar{\mathcal{S}}$ denotes the (Zariski) closure of $\mathcal{S}$. Such mappings will be referred to as invertible $\mathcal{S}$ preservers and $\bar{\mathcal{S}}$ preservers, respectively.

Some special cases of our problem have been considered by other authors. In [W], Watkins characterized invertible $\mathcal{S}(A)$ preservers for a diagonal matrix A with distinct eigenvalues. If A is a rank $n-1$ nilpotent matrix, then $\overline{\mathcal{S}(A)}$ is the set of all nilpotent matrices, and the invertible $\overline{\mathcal{S}(A)}$ preservers of such an A was characterized in [BPW]. In [H], Howard considered invertible linear operators that map the set of matrices annihilated by a given polynomial $f(x)$ with at least two distinct roots into itself. It is not hard to check that the set considered by Howard can be written as a finite union of similarity classes. Thus Howard's result is covered by ours. We also characterize those nonzero linear operators that commute with a given polynomial $f(x)$. This extends a result of Chan and Lim [CL1] on linear operators commuting with the function $f(x) = x^k$ for some given positive integer $k \geq 2$.

The first author was supported by NSF grant 91-00344 and a Faculty Research Grant from the College of William and Mary. This research was done while the author was visiting San Diego State University in November of 1992, during his sabbatical leave.

The second author was supported by NSF grant DMS 90-07048.

Received by the editors January 20, 1993; revised July 21, 1993.

AMS subject classification: 15A42.

https://doi.org/10.4153/CMB-1994-055-0 Published online by Cambridge University Press

We prove our main theorem in Section 2 and discuss some of its consequences in Section 3. In Section 4, we consider the possibility of removing the invertibility assumption in our results. Partial answers to a conjecture in [H] are given. In Section 5, we characterize those linear operators on M_n with $\mathbb{F} = \mathbb{C}$ that commute with an analytic function, and mention some extensions of our results to more general fields.

We shall use $\{E_{11}, E_{12}, \ldots, E_{nn}\}$ to denote the standard basis of M_n, and use M_n' to denote the collection of matrices in M_n with zero trace.

2. **Main theorem.** If $\mathcal{S}$ is a union of similarity classes of scalar matrices, then a linear operator ϕ on M_n satisfies $\phi(\mathcal{S}) \subseteq \mathcal{S}$ if and only if $\phi(I) = \mu I$ for some suitable $\mu \in \mathbb{F}$. For other cases, we have the following theorem.

THEOREM 2.1. *Let $A_1, \ldots, A_k \in M_n$ and $\mathcal{S} = \mathcal{S}(A_1) \cup \cdots \cup \mathcal{S}(A_k) \not\subseteq \{\lambda I : \lambda \in \mathbb{F}\}$. An invertible linear operator ϕ on M_n satisfies $\phi(\mathcal{S}) \subseteq \mathcal{S}$ or $\phi(\bar{\mathcal{S}}) \subseteq \bar{\mathcal{S}}$ if and only if one of the following conditions holds.*

(i) $\mathcal{S} \subseteq M_n'$ and ϕ is of the form

$$X \longmapsto (\operatorname{tr} X)D + \mu S^{-1}\big(X - (\operatorname{tr} X)I/n\big)S$$

or

$$X \longmapsto (\operatorname{tr} X)D + \mu S^{-1}\big(X^t - (\operatorname{tr} X)I/n\big)S,$$

for some $D \in M_n \setminus M_n'$, invertible $S \in M_n$, and $\mu \in \mathbb{F}$ such that $\big(\mu\mathcal{S}(A_1), \ldots, \mu\mathcal{S}(A_k)\big)$ is a permutation of $\big(\mathcal{S}(A_1), \ldots, \mathcal{S}(A_k)\big)$.

(ii) $\mathcal{S} \not\subseteq M_n'$ and ϕ is of the form

$$X \longmapsto \nu(\operatorname{tr} X)I/n + \mu S^{-1}\big(X - (\operatorname{tr} X)I/n\big)S$$

or

$$X \longmapsto \nu(\operatorname{tr} X)I/n + \mu S^{-1}\big(X^t - (\operatorname{tr} X)I/n\big)S,$$

for some invertible $S \in M_n$, and $\nu, \mu \in \mathbb{F}$ such that $\big(\mathcal{S}(\hat{A}_1), \ldots, \mathcal{S}(\hat{A}_k)\big)$ is a permutation of $\big(\mathcal{S}(A_1), \ldots, \mathcal{S}(A_k)\big)$, where $\hat{A}_i = \nu(\operatorname{tr} A_i)I/n + \mu\big(A_i - (\operatorname{tr} A_i)I/n\big)$ for $i = 1, \ldots, k$.

Furthermore, in conditions (i) and (ii), the element $\nu \in \mathbb{F}$ satisfies $\nu^p = 1$ for some positive integer p; and μ can be any nonzero element in $\mathbb{F}$ if all $\big(A_i - (\operatorname{tr} A_i)I/n\big)$'s are nilpotent matrices, otherwise $\mu^q = 1$ for some positive integer q.

To illustrate our theorem, we exhibit some examples of $\mathcal{S}$ whose invertible preservers are of the form (i) or (ii) in Theorem 2.1:

1. If $\mathcal{S} = \mathcal{S}(E_{12})$, then an invertible $\mathcal{S}$ preserver is of the form in (i), where μ can be any nonzero element in $\mathbb{F}$.
2. If $\mathcal{S} = \mathcal{S}(E_{12} + E_{21})$, then an invertible $\mathcal{S}$ preserver is of the form in (i) with $\mu = \pm 1$.
3. If $\mathcal{S} = \mathcal{S}(I + E_{12}) \cup \mathcal{S}(-I + E_{12})$, then an invertible $\mathcal{S}$ preserver is of the form in (ii) such that $\nu = \pm 1$ and μ can be any any nonzero element in $\mathbb{F}$.

4. If $\mathcal{S} = \mathcal{S}(I + E_{12} + E_{21}) \cup \mathcal{S}(-I + E_{12} + E_{12})$, then an invertible $\mathcal{S}$ preserver is of the form in (ii) with $\nu = \pm 1$ and $\mu = \pm 1$.

We establish several lemmas to prove Theorem 2.1. The key step in our proof is to reduce the general problem to the case when $\mathcal{S}$ is a collection of nilpotent matrices (*cf.* Lemma 2.3). Then we apply some elementary theory of linear algebraic groups to get the desired conclusion. We first state the following lemma (see [H, Lemma 1], [Hum, pp. 58–61]).

LEMMA 2.2. *Let $A \in M_n$. Then*

(a) $\overline{\mathcal{S}(A)}$ is an irreducible algebraic set in M_n, and the set of smooth points of $\overline{\mathcal{S}(A)}$ is $\mathcal{S}(A)$.

(b) The set of nonsmooth points of $\overline{\mathcal{S}(A)}$ is a finite union of other similarity orbits $\mathcal{S}(B_i)$.

LEMMA 2.3. *Let $\mathcal{S}$ satisfy the hypothesis of Theorem 2.1, and let $\mathcal{T}$ be the collection of $X \in M_n$ such that there exists $B \in \bar{\mathcal{S}}$ satisfying $B+\mu X \in \bar{\mathcal{S}}$ for all $\mu \in \mathbb{F}$. Then $\mathcal{T} \neq \{0\}$ is a union of similarity classes of nilpotent matrices. Moreover, if ϕ is an invertible linear operator on M_n mapping $\bar{\mathcal{S}}$ into itself, then ϕ maps $\mathcal{T}$ onto itself.*

PROOF. Suppose $A \in \mathcal{S}$ is nonscalar. Then A can be written as $R^{-1}(D + N)R$ for some diagonal matrix D and nonzero nilpotent matrix N. Clearly, $R^{-1}NR \in \mathcal{T}$ and hence $\mathcal{T} \neq \{0\}$.

Notice that for every B in $\bar{\mathcal{S}}$, the characteristic polynomial of B is the same as one of the A_i, and hence B has the same eigenvalues as A_i. If $X \in \mathcal{T}$ has a nonzero eigenvalue, then for any $B \in \bar{\mathcal{S}}$ there exists $\mu \in \mathbb{F}$ such that $B + \mu X$ does not have the same eigenvalues as A_j for all $j = 1, \ldots, k$. Thus X must be nilpotent. It is clear that if $X \in \mathcal{T}$ then $\mathcal{S}(X) \subseteq \mathcal{T}$. Thus the first assertion in the lemma follows.

Now suppose ϕ is an invertible linear operator on M_n mapping $\bar{\mathcal{S}}$ into itself. Since $\bar{\mathcal{S}}$ is an algebraic set, by the result of Dixon [D] we may assume that $\phi(\bar{\mathcal{S}}) = \bar{\mathcal{S}}$. It is now clear that if $X \in \mathcal{T}$ then both $Y = \phi(X)$ and $Z = \phi^{-1}(X)$ also belong to $\mathcal{T}$. Hence the result follows. ■

LEMMA 2.4. *Suppose ϕ is an invertible linear operator on M_n' mapping the set of rank one nilpotent matrices into itself. Then ϕ on M_n' is of the form $X \longmapsto \mu S^{-1}XS$ or $X \longmapsto \mu S^{-1}X^tS$ for some nonzero $\mu \in \mathbb{F}$ and some invertible $S \in M_n$.*

PROOF. Under our assumption, we can use the arguments in [BPW, pp. 43–45] to get the conclusion. ■

LEMMA 2.5. *Let $\mathcal{T}$ be a union of similarity classes of nilpotent matrices. Assume that $\mathcal{T} \neq 0$. If ϕ is an invertible linear operator on M_n' such that $\phi(\mathcal{T}) \subseteq \mathcal{T}$, then there exists an invertible $S \in M_n$ and a nonzero scalar $\mu \in \mathbb{F}$ such that ϕ on M_n' is of the form $A \longmapsto \mu S^{-1}AS$ or $A \longmapsto \mu S^{-1}A^tS$.*

PROOF. There is no harm in assuming that $\mathcal{T}$ is closed, otherwise replace $\mathcal{T}$ by $\bar{\mathcal{T}}$. We use induction on the dimension of $\mathcal{T}$. Let $\mathcal{T}'$ be the set of nonsmooth points of $\mathcal{T}$.

This is closed and is invariant under ϕ. Thus, we can replace $\mathcal{T}$ by $\mathcal{T}'$ unless $\mathcal{T}' = 0$. This occurs if and only if $\mathcal{T}$ consists of rank one matrices. Thus ϕ preserves rank one nilpotent matrices and the result follows by Lemma 2.4. ∎

PROOF OF THEOREM 2.1. The "if" part is clear. To prove the "only if" part, notice that $\phi(\mathcal{S}) \subseteq \mathcal{S}$ implies $\phi(\bar{\mathcal{S}}) \subseteq \bar{\mathcal{S}}$. Thus we may assume $\phi(\bar{\mathcal{S}}) \subseteq \bar{\mathcal{S}}$. Define $\mathcal{T}$ as in Lemma 2.3. Then ϕ maps $\mathcal{T}$ onto itself. By Lemma 2.5, ϕ on M_n' is of the form $X \longmapsto \mu S^{-1}XS$ or $X \longmapsto \mu S^{-1}X^tS$ for some invertible $S \in M_n$ and some nonzero $\mu \in \mathbb{F}$. For simplicity, we assume $\phi(X) = \mu X$ for all $X \in M_n'$. If not, consider $\hat{\phi}(X) = S\phi(X)S^{-1}$ or $S\phi(X^t)S^{-1}$ instead of ϕ.

Since ϕ is invertible and satisfies $\phi(\bar{\mathcal{S}}) \subseteq \bar{\mathcal{S}}$, we have $\phi(\bar{\mathcal{S}}) = \bar{\mathcal{S}}$ by a result of Dixon [D]. Notice that $\bar{\mathcal{S}} = \mathcal{S}(B_1)\cup\cdots\mathcal{S}(B_m)$ for some $B_1,\ldots,B_m \in M_n$ by Lemma 2.2. We conclude that $\Big(\phi\big(\mathcal{S}(B_1)\big),\ldots,\phi\big(\mathcal{S}(B_m)\big)\Big)$ is a permutation of $\big(\mathcal{S}(B_1),\ldots,\mathcal{S}(B_m)\big)$. In particular, $\Big(\phi\big(\mathcal{S}(A_1)\big),\ldots,\phi\big(\mathcal{S}(A_k)\big)\Big)$ is a permutation of $\big(\mathcal{S}(A_1),\ldots,\mathcal{S}(A_k)\big)$.

Suppose $\mathcal{S} \subseteq M_n'$. Let $D = \phi(I)/n$. If all A_i are nilpotent, then $\phi\big(\mathcal{S}(A_i)\big) = \mathcal{S}(\mu A_i)$ for any nonzero $\mu \in \mathbb{F}$. Suppose not all A_i are nilpotent, say A_1 is not nilpotent. Since $\mathcal{S}\big(\phi^r(A_1)\big) = \mathcal{S}(\mu^rA_1) \in \{\mathcal{S}(A_i) : i = 1,\ldots,k\}$ for all $r = 1,2,\ldots$, it follows that $\mathcal{S}(\mu^rA_1) = \mathcal{S}(\mu^sA_1)$ for some $1 \leq r < s$. Thus $\mu^{s-r}A_1$ and A_1 have the same eigenvalues that are not all equal to zero. Thus $\mu^q = 1$ for some positive integer q as asserted by the last statement of the theorem.

Now suppose $\mathcal{S} \not\subseteq M_n'$. We claim that $D = \phi(I)$ is a scalar matrix. To prove our claim, let $X \in \mathcal{S}$ with $\operatorname{tr} X \neq 0$. Notice that the maximal component in $\bar{\mathcal{S}}$ containing X must be of the form $\overline{\mathcal{S}(A)}$ for some $A \in \{A_1,\ldots,A_k\}$. Furthermore, since X and A have the same characteristic polynomial, $\operatorname{tr} A = \operatorname{tr} X \neq 0$. If $\phi\big(\overline{\mathcal{S}(A)}\big) = \overline{\mathcal{S}(A')}$, then ϕ will map the set of smooth points of $\overline{\mathcal{S}(A)}$ onto the set of smooth points of $\overline{\mathcal{S}(A')}$. If A is a scalar matrix, then $\overline{\mathcal{S}(A)}$ is a singleton. Thus $\overline{\mathcal{S}(A')}$ is also a singleton, and hence $\phi(A) = A'$ is a scalar matrix. Therefore our claim follows. Suppose A is not a scalar matrix. Since $R^{-1}AR$ is a smooth point of $\overline{\mathcal{S}(A)}$ for any invertible $R \in M_n$, it follows that $\phi(R^{-1}AR) = \phi(R^{-1}\hat{A}R) + (\operatorname{tr} A)\phi(I)/n = \mu(R^{-1}\hat{A}R) + (\operatorname{tr} A)D/n$ is a smooth point of $\overline{\mathcal{S}(A')}$, where $\hat{A} = A - (\operatorname{tr} A)I/n$. Thus $\mu(R^{-1}\hat{A}R) + (\operatorname{tr} A)D/n$ is similar to A' for all invertible $R \in M_n$. If $D = \phi(I)$ is not a scalar matrix, then there exists an invertible U such that $U^{-1}DU$ is in upper triangular form with its $(1,2)$ entry equal to one. Since $\hat{A}$ is not a scalar matrix, $\hat{A}$ is similar to an upper triangular matrix $\hat{A}_\eta$ such that the $(1,2)$ entry of $\hat{A}_\eta$ equals η for any given $\eta \in \mathbb{F}$. Let $R \in M_n$ (depending on η) be such that $U^{-1}R^{-1}\hat{A}RU = W\hat{A}_\eta W^{-1}$, where $W = E_{12} + E_{21} + \sum_{i=3}^n E_{ii}$. By a suitable choice of η, the spectrum of $\mu(R^{-1}\hat{A}R) + (\operatorname{tr} A)D/n$ will be different from the spectrum of A', which is a contradiction. Thus D must be a scalar matrix as we claimed.

Suppose $\phi(I) = \nu I$. Using arguments similar to the previous case, we can find an $A \in \bar{\mathcal{S}}$ with nonzero trace and show that $\phi^p\big(\mathcal{S}(A)\big) = \mathcal{S}\big(\nu^p(\operatorname{tr} A)I/n + \mu^p\hat{A}\big) = \mathcal{S}(A)$ for some positive integer p, where $\hat{A} = A - (\operatorname{tr} A)I/n$. Thus $\nu^p = 1$. If $A - (\operatorname{tr} A)I/n$ is nilpotent for any $A \in \bar{\mathcal{S}}$, then we are done. If there exists $A \in \bar{\mathcal{S}}$ such that $\hat{A} = A - (\operatorname{tr} A)I/n$ is not nilpotent, then one can use arguments similar to the previous case to show that

$\phi^p\big(\mathcal{S}(A)\big) = \mathcal{S}\big(\nu^p(\operatorname{tr} A)I/n + \mu^p\hat{A}\big) = \mathcal{S}(A)$ for some positive integer r. Hence $\mu^r\hat{A}$ and $\hat{A}$ have the same eigenvalues, and thus $\mu^q = 1$ for some positive integer q as asserted. ∎

3. **Related results.** In this section we discuss several consequences of Theorem 2.1. First of all, it covers a result due to Watkins [W]. There is a slight error in the statement of the theorem in [W]. By private communications with W. Watkins, we know that the statement and proof of the theorem in [W] can be easily modified to give the following result.

COROLLARY 3.1. *Suppose A is diagonalizable and has n distinct eigenvalues. An invertible linear operator ϕ on M_n satisfies $\phi\big(\mathcal{S}(A)\big) \subseteq \mathcal{S}(A)$ if and only if ϕ satisfies condition (i) or (ii) in Theorem 2.1.*

Notice that if A is a nilpotent matrix with rank $n-1$, then $\overline{\mathcal{S}(A)}$ is the set of all nilpotent matrices. Denote by M_n' the collection of matrices in M_n with trace zero. We have the following result which was proved in [BPW].

COROLLARY 3.2. *An invertible linear operator ϕ on M_n' maps the set of nilpotent matrices into itself if and only if ϕ on M_n' is of the form $X \longmapsto \mu S^{-1}XS$ or $X \longmapsto \mu S^{-1}X^tS$ for some nonzero $\mu \in \mathbb{F}$ and some invertible $S \in M_n$.*

As mentioned in the introduction, our main result covers a theorem of Howard [H] concerning the linear operators on M_n mapping the set of matrices A satisfying $f(A) = 0$ for a given polynomial $f(x)$ that has at least two distinct roots. In the following theorem, part (c.ii) is due to Howard [H] (see also [CL1, Theorem 3]), and part (c.i) deals with the case which is not treated in [H].

THEOREM 3.3. *Let f be a polynomial over $\mathbb{F}$ of degree at least two. Denote by $\mathcal{F}$ the set of all matrices X in M_n satisfying $f(X) = 0$. An invertible linear operator ϕ on M_n satisfies $\phi(\mathcal{F}) \subseteq \mathcal{F}$ if and only if one of the following holds.*

(a) $f(x) = x^k$ and ϕ is of form (i) in Theorem 2.1, where μ can be any nonzero element in $\mathbb{F}$.

(b) $f(x) = (x-\alpha)^k$ with $\alpha \neq 0$, and ϕ is of form (ii) in Theorem 2.1, where $\nu = 1$ and μ can be any nonzero element in $\mathbb{F}$.

(c.i) $n = 2$, f has at least two distinct roots, and ϕ is of the form (ii) in Theorem 2.1, where $\nu, \mu \in \mathbb{F}$ with $\nu^p = 1$ and $\mu^q = 1$ for some positive integers p and q such that $(x-\alpha)(x-\beta)$ is a factor of f if and only if $\big(x - \nu(\alpha+\beta)/2 + \mu(\alpha-\beta)/2\big) \cdot \big(x - \nu(\alpha+\beta)/2 - \mu(\alpha-\beta)/2\big)$ is a factor of f.

(c.ii) $n \geq 3$, f has at least two distinct roots, and ϕ is of the form (ii) in Theorem 2.1, where $\mu^q = 1$ such that $f(x) = x^r g(x^q)$ for some polynomial g.

PROOF. The "if" part can be readily verified. For the "only if" part, notice that $\mathcal{F} = \mathcal{S}(A_1) \cup \cdots \cup \mathcal{S}(A_k)$ for some $A_1, \ldots, A_k \in M_n$. By Theorem 2.1, we conclude that ϕ is of the form (i) or (ii) in Theorem 2.1.

Suppose $f(x) = x^k$. Then clearly, ϕ satisfies condition (i) with $D = \phi(I)/n$. Suppose $f(x) = (x-\alpha)^k$ with $\alpha \neq 0$. Then $A - (\operatorname{tr} A)I/n$ is nilpotent for all $A \in \mathcal{F}$. Thus ϕ is of the form (ii) in Theorem 2.1 for some $\nu, \mu \in \mathbb{F}$, where μ can be any nonzero element. Clearly, $\alpha I \in \mathcal{F}$ and $\phi(\alpha I) = \alpha I$. Thus $\nu = 1$.

Suppose f has at least two distinct roots. Then there exists $A \in \mathcal{F}$ with nonzero trace, and there exists $A' \in \mathcal{F}$ such that $A' - (\operatorname{tr} A')I/n$ is not nilpotent. Thus ϕ satisfies condition (ii) of Theorem 2.1 for some $\nu, \mu \in \mathbb{F}$ with $\nu^p = 1$ and $\mu^q = 1$ for some positive integers p and q. Since ϕ permutes the similarity classes $\mathcal{S}(A_1), \ldots, \mathcal{S}(A_k)$, we get condition (c.i) if $n = 2$. If $n \geq 3$, we claim that $\nu = \mu$. If it is not true, let $A = \operatorname{diag}(\alpha_1, \alpha_2, \ldots, \alpha_2) \in \mathcal{F}$ be such that $\alpha_1 \neq \alpha_2$. Notice that all $\phi(A), \phi^2(A), \ldots \in \mathcal{F}$ are of the form $\operatorname{diag}(\beta_1, \beta_2, \ldots, \beta_2)$ with $\beta_1 \neq \beta_2$. Since there are finitely many matrices in $\mathcal{F}$ of the above form, there exists a positive integer s such that $\phi^s(A) = A$. Now by the arguments in [H, 175–176] one can show that μ/ν and $(1 - \mu/\nu)/n$ are both roots of unity to get a contradiction. Thus we must have $\mu = \nu$ as asserted. ∎

It would be nice to have a simple description for the condition on the polynomial f in terms of p and q in (c.i). Next we study those linear operators ϕ on M_n that satisfy $\phi\big(f(X)\big) = f\big(\phi(X)\big)$ for all $X \in M_n$. For $f(x) = x^k$ with $k \geq 2$ this problem was studied in [CL1, Theorem] without the invertibility assumption on ϕ (see Section 4).

THEOREM 3.4. *Let f be a polynomial over $\mathbb{F}$ of degree at least two. An invertible linear operator ϕ on M_n satisfies $\phi\big(f(X)\big) = f\big(\phi(X)\big)$ for all X if and only if ϕ is of the form (ii) in Theorem 2.1 with $\nu = \mu$ satisfying $\mu^k = 1$ such that $f(x) = xg(x^k)$ for some polynomial g.*

PROOF. The "if" part can be easily verified. For the only if part, let $f_\eta(x) = f(x) - \eta x$ for any $\eta \in \mathbb{F}$ and let $\mathcal{F}_\eta$ be the collection of $X \in M_n$ satisfying $f_\eta(X) = 0$. If ϕ commutes with f and if $X \in \mathcal{F}_\eta$, then $f_\eta\big(\phi(X)\big) = f\big(\phi(X)\big) - \eta\phi(X) = \phi\big(f(X)\big) - \phi(\eta X) = \phi\big(f_\eta(X)\big) = 0$. Thus $\phi(\mathcal{F}_\eta) \subseteq \mathcal{F}_\eta$ for all $\eta \in \mathbb{F}$. In particular, we can choose $\eta \in \mathbb{F}$ such that $f_\eta(x)$ is not of the form $(x - \alpha)^k$ for any $\alpha \in \mathbb{F}$. Thus ϕ satisfies condition (i) or (ii) of Theorem 2.1. Furthermore, suppose ϕ is of the form $X \longmapsto \mu S^{-1}XS$ or $X \longmapsto \mu S^{-1}X^tS$ such that $\mu \neq 1$ is a k-th root of unity. Then $f(x) = x^r g(x^k)$ for some polynomial k. It follows that $f_\eta(x) = x^r g(x) - \eta x$ is of the form $x^s \hat{g}(x^k)$ for all $\eta \in \mathbb{F}$. Thus k must be a factor of $r - 1$ and the result follows. ∎

4. **Removal of invertibility assumption.**

In this section we consider the possibility of removing the invertibility assumption in our results obtained in the previous sections. First we consider the following example showing that the invertibility assumption is necessary in general.

EXAMPLE. Suppose A is nonscalar and $\operatorname{tr} A \neq 0$. Then one can construct many singular $\mathcal{S}(A)$ and $\overline{\mathcal{S}(A)}$ preservers. For instance, if $B = (b_{ij}) \in \mathcal{S}(A)$ is in upper triangular Jordan form, one can define ϕ by

$$X \longmapsto (\operatorname{tr} X/\operatorname{tr} A)B + \sum_{i<j,\, b_{ii} \neq b_{jj}} F_{ij}(X)E_{ij}$$

for any linear functionals F_{ij} on M_n. One easily checks that such a ϕ is singular and preserves $\mathcal{S}(A)$ and $\overline{\mathcal{S}(A)}$.

Nonetheless, we conjecture that:

I. Suppose $\mathcal{S}$ satisfies the hypotheses of Theorem 2.1. If ϕ is a linear operator on M_n with $\operatorname{rank}(\phi) > n(n-1)/2 + 1$ such that $\phi(\mathcal{S}) \subseteq \mathcal{S}$ or $\phi(\bar{\mathcal{S}}) \subseteq \bar{\mathcal{S}}$, then ϕ is invertible.

If we know more about the matrices $A_1, \ldots, A_k$ in the definition of $\mathcal{S}$, we might be able to reduce the lower bound of the rank requirement and prove that $\mathcal{S}$ or $\bar{\mathcal{S}}$ preservers are invertible. In particular, we believe that:

II. If $A \in M_n'$ is nonzero, then all $\mathcal{S}(A)$ preservers are invertible on M_n'.

III. If $A \in M_n'$ is not nilpotent, then all $\overline{\mathcal{S}(A)}$ preservers are invertible on M_n'.

The following result gives some support to Conjectures II and III.

THEOREM 4.1. *Suppose $A \in M_n'$ is a nonderogatory matrix and not nilpotent. If ϕ is a linear operator on M_n' satisfying $\phi\big(\mathcal{S}(A)\big) \subseteq \mathcal{S}(A)$ or $\phi\big(\overline{\mathcal{S}(A)}\big) \subseteq \overline{\mathcal{S}(A)}$, then ϕ is invertible.*

PROOF. Obviously if $\phi\big(\mathcal{S}(A)\big) \subseteq \mathcal{S}(A)$, then $\phi\big(\overline{\mathcal{S}(A)}\big) \subseteq \overline{\mathcal{S}(A)}$, so we assume condition $\phi\big(\overline{\mathcal{S}(A)}\big) \subseteq \overline{\mathcal{S}(A)}$. Since A is non-derogatory, one easily checks that the set $\mathcal{T}$ defined in Lemma 2.3 is the set of all nilpotent matrices in M_n'. Furthermore, if $\phi\big(\overline{\mathcal{S}(A)}\big) \subseteq \overline{\mathcal{S}(A)}$ then $\phi(\mathcal{T}) \subseteq \mathcal{T}$.

Now suppose $B \in M_n'$ is nonzero and $\phi(B) = 0$. Then (*e.g.*, see [JS]) there exists $R \in M_n$ such that $R^{-1}BR$ has diagonal entries equal to the eigenvalues of A. Let $R^{-1}BR = U + L$, where U is upper triangular and L is strictly lower triangular. Then $N = RLR^{-1}$ is nilpotent and $B + N \in \overline{\mathcal{S}(A)}$. But then $\phi(B+N) = \phi(N)$ is nilpotent and belongs to $\overline{\mathcal{S}(A)}$. Since $\overline{\mathcal{S}(A)}$ has no nilpotent matrices, this contradiction concludes the proof. ∎

Notice that if A is a non-derogatory nilpotent matrix, *i.e.*, A is a rank $n-1$ nilpotent matrix, then $\overline{\mathcal{S}(A)}$ is the set of all nilpotent matrices. In such case, there are examples of singular $\overline{\mathcal{S}(A)}$ preservers (*e.g.*, see [BPW]). However, in view of Conjecture II, we believe that all $\mathcal{S}(A)$ preservers are invertible. In fact, if A is a rank one nilpotent matrix, then $\mathcal{S}(A)$ is the set of all rank one nilpotent matrices. By the next result, we see that for such an A, any $\mathcal{S}(A)$ preserver is invertible.

THEOREM 4.2. *Let $1 \leq k < n$ and let $\mathcal{E}_k$ be the set of all nonzero nilpotent matrices with rank at most k. A linear operator ϕ on M_n' satisfies $\phi(\mathcal{E}_k) \subseteq \mathcal{E}_k$ if and only if ϕ is of the form $X \longmapsto \mu S^{-1}XS$ or $X \longmapsto \mu S^{-1}X^tS$ for some nonzero $\mu \in \mathbb{F}$ and some invertible $S \in M_n$.*

PROOF. The "if" part is clear. For the converse, notice that $Z = \mathcal{E}_k \cup \{0\}$ is a homogeneous algebraic variety. By the assumption on ϕ and by Lemma 3 in [CL2], we conclude that ϕ is invertible on M_n'. The result then follows from Theorem 2.1. ∎

Suppose f is a polynomial over $\mathbb{F}$ with at least two distinct roots. Denote by $\mathcal{F}$ the set of $X \in M_n$ satisfying $f(X) = 0$. Howard in [H] conjectured that:

IV. If $f(0) \neq 0$, then a linear operator ϕ on M_n with $n \geq 3$ satisfying $\phi(\mathcal{F}) \subseteq \mathcal{F}$ is invertible.

Clearly, if x^r divides $f(x)$, then ϕ defined by

$$\phi(X) = \sum_{r\geq i>j\geq 1} F_{ij}(X)E_{ij}$$

for some linear functionals F_{ij} is a singular linear operator that satisfies $\phi(\mathcal{F}) \subseteq \mathcal{F}$. Nevertheless, we believe that:

V. If $f(x) = x^r g(x)$ is such that $g(0) \neq 0$ and if $\phi(\mathcal{F}) \subseteq \mathcal{F}$, then either $\phi(M_n) \subseteq \mathcal{F}$ or ϕ is invertible.

The following theorem gives some support to Conjecture IV.

THEOREM 4.3. *Let f be a polynomial over $\mathbf{F}$ of degree at least two such that $f(0) \neq 0$. Denote by $\mathcal{F}$ the set of matrices $X \in M_n$ satisfying $f(X) = 0$. Suppose there exists a nonderogatory $A \in \mathcal{F}$ with $\operatorname{tr} A = 0$. If ϕ is a linear operator on M_n such that $\phi(\mathcal{F}) \subseteq \mathcal{F}$, then ϕ is invertible unless $n = 2$ and $f(x) = x^2 - \alpha^2$ for some nonzero α.*

PROOF. Suppose f, $\mathcal{F}$, A and ϕ satisfy the hypotheses of the theorem. Since $f(0) \neq 0$ and $f(A) = 0$, where $\operatorname{tr} A = 0$, A has at least two distinct eigenvalues. Hence f has at least two distinct roots.

First, we show that $\phi(M_n') \subseteq M_n'$. Suppose N is a nilpotent matrix. Then there exists $\hat{A}$ similar to A such that $\hat{A} + \mu N \in \overline{\mathcal{S}(A)} \subseteq \mathcal{F}$ for all $\mu \in \mathbf{F}$. As a result, if $\mathcal{T}$ is defined as in Lemma 2.3, then $\mathcal{T}$ is the set of all nilpotent matrices. Since $\phi(\mathcal{T}) \subseteq \mathcal{T}$ and the span of $\mathcal{T}$ is M_n', we get the desired conclusion.

Second, we show that ϕ is invertible on M_n'. If it is not true, then there exists a nonzero $B \in M_n'$ such that $\phi(B) = 0$. Using the arguments in the proof of Theorem 4.1, we can find a nilpotent matrix N such that $B + N \in \overline{\mathcal{S}(A)} \subseteq \mathcal{F}$. It follows that $\phi(B + N) = \phi(N)$ is a nilpotent matrix and does not belong to $\mathcal{F}$ as $f(0) \neq 0$. Thus our assertion is true.

To complete our proof, we show that $\phi(I) \notin M_n'$, and hence ϕ is surjective.

Since ϕ is invertible on M_n' and maps nilpotent matrices to nilpotent matrices, by Corollary 3.2, ϕ is of the form $X \mapsto \mu S^{-1}XS$ or $X \mapsto \mu S^{-1}X^tS$ for all $X \in M_n'$. For simplicity, we assume $\phi(X) = \mu X$ for all $X \in M_n'$, otherwise replace ϕ by $\hat{\phi}$ defined by $\hat{\phi}(X) = S\phi(X)S^{-1}$ or $\hat{\phi}(X) = S\phi(X^t)S^{-1}$.

Suppose $\phi(I) = Y \in M_n'$. Then $Y \neq 0$, otherwise $\phi(\alpha I) = 0 \notin \mathcal{F}$ for any α satisfying $f(\alpha) = 0$, *i.e.*, $\alpha I \in \mathcal{F}$, which is a contradiction. Since Y is not a scalar matrix, there exists $R \in M_n$ such that $R^{-1}YR$ is in upper triangular form with $(1, 2)$ entry equal to one.

We consider two cases.

CASE 1. There are two distinct roots α, β of f such that $\alpha + (n-1)\beta \neq 0$. Let $C \in \mathcal{F}$ be such that $R^{-1}CR = \alpha E_{11} + \beta(E_{22} + \cdots E_{nn}) + \eta E_{21}$. By a suitable choice of η, the matrix $\phi(C) = \phi\big(C - (\operatorname{tr} C)I/n\big) + (\operatorname{tr} C)\phi(I)/n = \mu\big(C - (\operatorname{tr} C)I/n\big) + (\operatorname{tr} C)Y/n$ has an eigenvalue which is not a root of f, and hence $\phi(C) \notin \mathcal{F}$, which is a contradiction.

CASE 2. If Case 1 does not hold, then $n = 2$ and f only has roots α and $-\alpha$. If $f(x) \neq x^2 - \alpha^2$, we may assume that $\alpha I + E_{12} \in \mathcal{F}$. Furthermore, we may assume that $R^{-1}YR = E_{11} - E_{22}$. Let $C = R^{-1}(\eta(E_{11} - E_{22}) + E_{12} - \eta^2 E_{21})R$. Then $X = \alpha I + C \in \mathcal{F}$, but η can be chosen so that $\phi(X)$ has eigenvalues not equal to $\pm\alpha$. Thus $\phi(X) \notin \mathcal{F}$, which is a contradiction. Combining the two cases, we conclude that $\operatorname{tr}\phi(I) \neq 0$ as asserted. ■

Note that if $n = 2$ and $f(x) = x^2 - \alpha^2$ for some nonzero α, then $\phi(X) = (\operatorname{tr} X)(E_{11} - E_{22})/2 + \big(X - (\operatorname{tr} X)I/2\big)$ is a singular linear operator on M_2 satisfying $\phi(\mathcal{F}) \subset \mathcal{F}$.

Finally, consider those linear operators that commute with a given polynomial f. We have the following theorem.

THEOREM 4.4. *Let f be a polynomial over $\mathbb{F}$ of degree at least two. If ϕ is a linear operator on M_n satisfying $\phi\big(f(X)\big) = f\big(\phi(X)\big)$ for all $X \in M_n$, then either (i) ϕ is invertible, or (ii) $f(0) = 0$ and ϕ is the zero map.*

PROOF. We may assume that $f(x)$ is a monic polynomial of degree k with $k > 1$. By the assumption, for any $X \in M_n$ and any $\lambda \in \mathbb{F}$ we have $f\big(\phi(\lambda X)\big) = \phi\big(f(\lambda X)\big)$. Comparing the coefficients of λ^k on both sides, we conclude that $\phi(X^k) = \phi(X)^k$ for all $X \in M_n$. By the results in [CL1], either ϕ is invertible or ϕ is the zero map. Suppose $f(0) \neq 0$. Comparing the constant terms, *i.e.*, the coefficients of λ^0, on both sides of $f\big(\phi(\lambda X)\big) = \phi\big(f(\lambda X)\big)$, we conclude that $\phi(I) = I$ and hence ϕ cannot be the zero map. The result follows. ■

For other results supporting our conjectures, see [BrS], [BeP] and [CL1].

5. **Other extensions.** One can use the idea in the proof of Theorem 4.4 to study linear operators on M_n with $\mathbb{F} = \mathbb{C}$ that commute with an analytic function. Notice that if $f(z) = \sum_{i=0}^{\infty} a_i z^i$ is an analytic function on $\mathbb{C}$ and if ϕ is a linear operator on M_n commuting with f, then for any ε and any $A \in M_n$, we have

$$\sum_i a_i \varepsilon^i \phi(A)^i = f\big(\phi(\varepsilon A)\big) = \phi\big(f(\varepsilon A)\big) = \sum_i a_i \varepsilon^i \phi(A^i).$$

Thus ϕ commutes with the functions $f_i(z) = z^i$ whenever $a_i \neq 0$. Consequently, we have

THEOREM 5.1. *Suppose f is an analytic function on $\mathbb{C}$. A nonzero linear operator ϕ on M_n with $\mathbb{F} = \mathbb{C}$ commutes with f if and only if ϕ is of the form $X \longmapsto \mu S^{-1}XS$ or $X \longmapsto \mu S^{-1}X^tS$ for some invertible $S \in M_n(\mathbb{C})$ and some $\mu \in \mathbb{C}$ such that $\mu f(z) = f(\mu z)$ for all $z \in \mathbb{C}$.*

It has been pointed out by M.H. Lim that one may remove the assumption that $\mathbb{F}$ is algebraically closed in Theorem 2.1 to get the same conclusion. R. Guralnick observed this independently and obtained the detailed proof of Theorem 2.1 for fields of arbitrary characteristics. M.H. Lim also extended Theorem 3.3 to any infinite field in which f splits.

ACKNOWLEDGEMENT. The authors thank D. Ž. Đoković, R. Guralnick, M. H. Lim and W. Watkins for some helpful discussions. Thanks are also due to the referee for some valuable suggestions.

References

[BeP] L. B. Beasley and N. B. Pullman, *Linear operators preserving idempotent matrices over fields*, Linear Algebra Appl. **146**(1991), 7–20.

[BPW] E. P. Botta, S. Pierce and W. Watkins, *Linear transformations that preserve the nilpotent matrices*, Pacific J. Math. **104**(1983), 39–46.

[BrS] M. Brešar and P. Semrl, *Linear transformations preserving potent matrices*, Proc. Amer. Math. Soc., to appear.

[CL1] G. H. Chan and M. H. Lim, *Linear preservers on powers of matrices*, Linear Algebra Appl. **162-164**(1992), 615–626.

[CL2] ———, *Linear transformations on symmetric matrices II*, Linear and Multilinear Algebra **32**(1992), 319–326.

[Di] J. Dixon, *Rigid embeddings of simple groups in the general linear group*, Canad. J. Math. **29**(1977), 384–391.

[H] R. Howard, *Linear maps that preserve matrices annihilated by a polynomial*, Linear Algebra Appl. **30** (1980), 167–176.

[Hum] J. E. Humphreys, *Linear Algebraic Groups*, Graduate Texts in Math. **21**, Springer, New York, 1975.

[JS] C. R. Johnson and H. Shapiro, *Mathematical aspects of the relative gain array* $[A \cdot A^{-1}]^*$, SIAM J. Algebra Discrete Math. **7**(1986), 627–644.

[W] W. Watkins, *Linear transformations that preserve a similarity class of matrices*, Linear and Multilinear Algebra **11**(1982), 19–22.

Department of Mathematics
The College of William and Mary
Williamsburg, Virginia 23187
U.S.A.

Department of Mathematical Sciences
San Diego State University
San Diego, California 92182
U.S.A.