Canad. Math. Bull. Vol. 63 (1), 2020 pp. 94–105
http://dx.doi.org/10.4153/S0008439519000250

Triangularization of Matrices and Polynomial Maps

Yueyue Li, Yan Tian, and Xiankun Du

Abstract. We present conditions for a set of matrices satisfying a permutation identity to be simultaneously triangularizable. As applications of our results, we generalize Radjavi's result on triangularization of matrices with permutable trace and results by Yan and Tang on linear triangularization of polynomial maps.

1 Introduction

Let K be a field. A set S of $n \times n$ matrices over K is called triangularizable if there exists an invertible matrix P over K such that $P^{-1}AP$ is an upper triangular matrix for all $A \in S$.

Triangularization of a set of matrices has been studied extensively in matrix theory [14]. Among others, Radjavi proved that a set S of $n \times n$ matrices is triangularizable if $\operatorname{tr}((AB - BA)B_1 \cdots B_s) = 0$ for all positive integers s and for all $A, B, B_1, \dots, B_s \in S$, where K is a field of characteristic either zero or greater than $n/2$ that contains the eigenvalues of every matrix in S [13, Theorem 1]. The Radjavi theorem generalizes several classic results on simultaneous triangularization of matrices, due to Levitzki, McCoy, Kaplansky, and others [14]. The Radjavi theorem also was generalized in [5,8].

On the other hand, linear triangularization of polynomial maps has been studied by several authors in the field of affine algebraic geometry [2–4, 9–11, 15, 17, 19, 21], because it relates to two famous open problems: the tame generators problem and the Jacobian conjecture [1,18]. Recently, Yan and Tang [21] proved that a polynomial map $F = X + H$ with the Jacobian matrix of H nilpotent is linearly triangularizable if there exists an integer $r \geq 2$ such that

$$JH(X^{(1)})JH(X^{(2)})\cdots JH(X^{(r)}) = JH(X^{(\sigma(1))})JH(X^{(\sigma(2))})\cdots JH(X^{(\sigma(r))})$$

for all permutation σ on $\{1, 2, \dots, r\}$.

Our aim is to generalize the Radjavi theorem and the result of Yan and Tang by investigating triangularizability of a σ-permutable set S of matrices, that is, S satisfies a permutation identity:

$$A_1A_2\cdots A_r = A_{\sigma(1)}A_{\sigma(2)}\cdots A_{\sigma(r)}, \text{ for all } A_1, A_2, \dots, A_r \in S,$$

Received by the editors September 20, 2018; revised March 10, 2019.
Published online on Cambridge Core September 18, 2019.
Supported by NSF of China (No. 11771176).
Corresponding author: Xiankun Du.
AMS subject classification: **15A21**, 14R15.
Keywords: matrix, Jacobian matrix, permutation, polynomial map, simultaneous triangularization.

for some nonidentity permutation σ. This paper is organized as follows. Section 2 is devoted to the permutability group $G(S)$ of a permutable subset S of a semigroup. It is proved that if S is permutable, then $G(S)$ contains 3-cycles $(k\ m\ n)$ for some positive integer d and for all distinct positive integers k, m, n with $k \equiv m \equiv n \mod d$. In Section 3 it is proved that a σ-permutable set S of $n \times n$ matrices with $\Delta(\sigma) = 1$ is triangularizable, where $\Delta(\sigma) = \gcd\{|\sigma(i) - i| \mid 1 \le i \le r\}$. This result is used to prove a generalization of the Radjavi theorem. In Section 4, we are concerned with triangularization of a permutable subspace of nilpotent matrices. In Section 5, as an application of our results, we generalize the result of Yan and Tang [21] on triangularization of polynomial maps.

Throughout this paper, let $\mathbb{N}$ be the set of positive integers. Denote by e_{ij} the $n \times n$ matrix unit, *i.e.*, the matrix with 1 in the (i, j) entry and 0 elsewhere, for $1 \le i, j \le n$.

2 Permutability Groups

In this section, let S denote a nonempty subset of a semigroup. For $k \in \mathbb{N}$, let $S^k = \{a_1 a_2 \cdots a_k \mid a_1, a_2, \dots, a_k \in S\}$ and let S^0 be empty.

Let S_n denote the symmetric group on the set $\{1, 2, \dots, n\}$ for $n \in \mathbb{N}$. The product of permutations is read from right to left.

Given a permutation $\sigma \in S_n$, S is called σ-permutable if

$$a_1 a_2 \cdots a_n = a_{\sigma(1)} a_{\sigma(2)} \cdots a_{\sigma(n)}$$

for all $a_1, a_2, \dots, a_n \in S$. If S is σ-permutable for some nonidentity permutation σ, then S is called permutable.

A permutation σ in S_m can naturally be considered as a permutation σ' in S_n for $m \le n$. However, S is not necessarily σ-permutable, even if S is σ'-permutable. For example, let $S = \{e_{11}, e_{12}\}$. Then S is $(1\,2)$-permutable for the 2-cycle $(1\,2) \in S_3$, but not for the 2-cycle $(1\,2) \in S_2$. To avoid this issue we will consider finitary permutations on $\mathbb{N}$.

Denote by $\mathrm{FSym}(\mathbb{N})$ the finitary symmetric group on $\mathbb{N}$ and consider S_n as a subgroup in the natural way for all $n \in \mathbb{N}$. Then $\mathrm{FSym}(\mathbb{N})$ is the union of all S_n, $n \in \mathbb{N}$. For $\sigma \in \mathrm{FSym}(\mathbb{N})$, the support of σ is defined by

$$\mathrm{supp}(\sigma) = \{i \in \mathbb{N} \mid \sigma(i) \neq i\}.$$

See [6] for more details on $\mathrm{FSym}(\mathbb{N})$.

Denote by $G(S)$ the set of $\sigma \in \mathrm{FSym}(\mathbb{N})$ satisfying

$$u a_1 a_2 \cdots a_n v = u a_{\sigma(1)} a_{\sigma(2)} \cdots a_{\sigma(n)} v$$

for some integer n greater than or equal to the maximum in $\mathrm{supp}(\sigma)$, for all sufficiently large $t \in \mathbb{N}$ and for all $u, v \in S^t$ and $a_1, a_2, \dots, a_n \in S$.

Remark 2.1 The set $G(S)$ is a subgroup of $\mathrm{FSym}(\mathbb{N})$, and a set S is permutable if and only if $G(S)$ is nontrivial.

https://doi.org/10.4153/S0008439519000250 Published online by Cambridge University Press

Remark 2.2 Under the canonical embedding of S_n into $\mathrm{FSym}(\mathbb{N})$, if S is σ-permutable for $\sigma \in S_n$, then $\sigma \in G(S)$, and conversely, if $\sigma \in G(S)$, then there exists $n \in \mathbb{N}$ such that $\sigma \in S_n$ and S is σ-permutable.

If S is permutable, let

$$d(S) = \min\left\{ |\sigma(i) - i| \;\middle|\; \sigma \in G(S) \text{ and } i \in \mathrm{supp}(\sigma) \right\}.$$

For $\sigma \in \mathrm{FSym}(\mathbb{N})$ and an integer $k \geq 0$, we define $\sigma_{[k]} \in \mathrm{FSym}(\mathbb{N})$ as follows

$$\sigma_{[k]}(i) = \begin{cases} i & \text{if } 1 \leq i \leq k, \\ \sigma(i-k) + k & \text{if } i > k, \end{cases}$$

that is,

$$\sigma_{[k]} = \begin{pmatrix} 1\,2 \cdots k & k+1 & k+2 & \cdots \\ 1\,2 \cdots k & k+\sigma(1) & k+\sigma(2) & \cdots \end{pmatrix}.$$

Lemma 2.3 *Let* $\sigma \in \mathrm{FSym}(\mathbb{N})$.

(i) *If* $\sigma \in G(S)$, *then* $\sigma_{[k]} \in G(S)$ *for all* $k \in \mathbb{N}$.
(ii) *If* $\sigma_{[k]} \in G(S)$, *for some* $k \in \mathbb{N}$, *then* $\sigma \in G(S)$.

Proof The proof is straightforward. ■

Lemma 2.4 *If S is permutable, then* $d(S) \mid \sigma(i) - \tau(i)$ *for all* $i \in \mathbb{N}$ *and all* $\sigma, \tau \in G(S)$.

Proof Let $T = \{\sigma(i) - i \mid \sigma \in G(S) \text{ and } i \in \mathbb{N}\}$. Then

$$\sigma(i) - \tau(i) = \sigma\tau^{-1}(\tau(i)) - \tau(i) \in T,$$

for all $\sigma, \tau \in G(S)$ and $i \in \mathbb{N}$. Thus, $T = \{\sigma(i) - \tau(i) \mid \sigma, \tau \in G(S) \text{ and } i \in \mathbb{N}\}$. Now we need only prove that T is a subgroup of the additive group of integers. Let $m = \sigma(i) - i$ and $n = \tau(j) - j$ for $\sigma, \tau \in G(S)$ and $i, j \in \mathbb{N}$. Since $\sigma_{[j]}(i+j) = \sigma(i) + j$ and $\tau_{[i]}(i+j) = \tau(j) + i$, by Lemma 2.3 we have

$$m - n = (\sigma(i) + j) - (\tau(j) + i) = \sigma_{[j]}(i+j) - \tau_{[i]}(i+j) \in T.$$

Thus T is an additive subgroup of the integers, as desired. ■

The following lemma is a well-known fact [6, Exercise 1.6.7]. We include here a proof for completeness.

Lemma 2.5 *Let G be a subgroup of* $\mathrm{FSym}(\mathbb{N})$. *If* $\mathrm{supp}(\sigma) \cap \mathrm{supp}(\tau) = \{m\}$, *for* $\sigma, \tau \in G$, *then the* 3*-cycle* $(\sigma^u(m)\ \ m\ \ \tau^v(m)) \in G$ *for all* $u, v \in \{-1, 1\}$.

Proof It is a routine matter to verify that $\tau\sigma\tau^{-1}\sigma^{-1} = (\sigma(m)\ \ m\ \ \tau(m))$. Thus $(\sigma(m)\ \ m\ \ \tau(m)) \in G$. Noting that $\mathrm{supp}(\sigma^u) \cap \mathrm{supp}(\tau^v) = \{m\}$, we see that $(\sigma^u(m)\ \ m\ \ \tau^v(m)) \in G$ for all $u, v \in \{-1, 1\}$. ■

Lemma 2.6 *If S is permutable and* $d(S) = d$, *then* $(1\ \ 1{+}d\ \ 1{+}2d) \in G(S)$.

Proof Let $d = \sigma(i_1) - i_1 > 0$ for some $\sigma \in G(S)$ and some $i_1 \in \mathbb{N}$. Let i_0 and i_s be the minimum and the maximum in $\mathrm{supp}(\sigma)$, respectively. Then $\mathrm{supp}(\sigma) \cap \mathrm{supp}(\sigma_{[i_s-i_0]}) = i_s$. By Lemma 2.5,

$$(\sigma(i_s)\ \ i_s\ \ \sigma_{[i_s-i_0]}(i_s)) = (\sigma(i_s)\ \ i_s\ \ \sigma(i_0)+i_s-i_0) \in G(S).$$

Since $i_0 < \sigma(i_0)$ by the minimality of i_0 and $\sigma(i_s) < i_s$ by the maximality of i_s, we have $\sigma(i_0) - i_0 = m_1 d$ and $i_s - \sigma(i_s) = m_2 d$ for some $m_1, m_2 \in \mathbb{N}$ by Lemma 2.4, whence $(i_s-m_2d\ \ i_s\ \ i_s+m_1d) \in G(S)$. It follows from Lemma 2.3 that $(i_s i_s + m_2 d i_s + (m_1 + m_2)d) \in G(S)$. Thus Lemma 2.3 implies that

$$(i_s+(m_1+m_2)d\ \ i_s+(m_1+2m_2)d\ \ i_s+2(m_1+m_2)d) \in G(S).$$

Applying Lemma 2.5 to the last two 3-cycles gives

$$(i_s\ \ i_s+(m_1+m_2)d\ \ i_s+2(m_1+m_2)d) \in G(S).$$

Let $n = m_1 + m_2$. Then $(i_s\ \ i_s+nd\ \ i_s+2nd) \in G(S)$. Now a simple induction gives $(i_s\ \ i_s+2^{k-1}nd\ \ i_s+2^k nd) \in G(S)$ for all $k \in \mathbb{N}$. For, if

$$(i_s\ \ i_s+2^{k-1}nd\ \ i_s+2^k nd) \in G(S),$$

then $(i_s+2^k nd\ \ i_s+(2^{k-1}+2^k)nd\ \ i_s+2^{k+1}nd) \in G(S)$ by Lemma 2.3, and so applying Lemma 2.5 to the last two 3-cycles yields $(i_s\ \ i_s+2^k nd\ \ i_s+2^{k+1}nd) \in G(S)$.

Take $N = 2^{k-1}nd$ for some $k \in \mathbb{N}$ such that $N > \max\{i_s - i_1, 2d\}$ and let $\tau = (i_s\ \ i_s+N\ \ i_s+2N)$. Then we see that $\mathrm{supp}(\sigma_{[i_s-i_1]}) \cap \mathrm{supp}(\tau) = \{i_s\}$, and so by Lemma 2.5 we have that

$$(i_s\ \ \sigma_{[i_s-i_1]}(i_s)\ \ \tau(i_s)) = (i_s\ \ i_s+d\ \ i_s+N) \in G(S).$$

Also $(i_s+d\ \ i_s+2d\ \ i_s+N+d) \in G(S)$ by Lemma 2.3. Since $N > 2d$, by Lemma 2.5, we have $(i_s\ \ i_s+d\ \ i_s+2d) \in G(S)$. Let $\rho = (1\ \ 1+d\ \ 1+2d)$. Then $\rho_{[i_s-1]} = (i_s\ \ i_s+d\ \ i_s+2d) \in G(S)$, and so $(1\ \ 1+d\ \ 1+2d) \in G(S)$ by Lemma 2.3. ■

Lemma 2.7 *If $d(S) = d$, then $G(S)$ contains the 3-cycles $(r\ s\ t)$ for all $r, s, t \in \mathbb{N}$ with $r \equiv s \equiv t \mod d$.*

Proof It suffices to prove that $(1 + kd\ \ 1 + md\ \ 1 + nd) \in G(S)$ for all integers k, m, n with $0 \leq k < m < n$, since in this case $(i + kd\ \ i + md\ \ i + nd) \in G(S)$ for $i \in \mathbb{N}$ by Lemma 2.3.

We first claim that $(1\ \ 1 + d\ \ 1 + kd) \in G(S)$ for all integers $k > 1$. Indeed, it follows by induction. By Lemma 2.6, $(1\ \ 1 + d\ \ 1 + 2d) \in G(S)$. Suppose $(1\ \ 1 + d\ \ 1 + (k-1)d) \in G(S)$ for $k > 2$. Then $(1 + d\ \ 1 + 2d\ \ 1 + kd) \in G$ by Lemma 2.3. Thus $(1\ \ 1 + d\ \ 1 + kd) \in G(S)$ by Lemma 2.5, as desired.

We next claim that $(1\ \ 1 + kd\ \ 1 + md) \in G(S)$ for all integers $0 < k < m$. If $k = 1$, then there is nothing to prove by the first claim. Suppose that $k > 1$. Then the first claim gives $\sigma = (1\ \ 1 + d\ \ 1 + kd) \in G(S)$ and $(1\ \ 1 + d\ \ 1 + md) \in G(S)$. Thus $\sigma^{-1}(1\ \ 1 + d\ \ 1 + md)\sigma = (\sigma^{-1}(1)\ \ \sigma^{-1}(1 + d)\ \ \sigma^{-1}(1 + md)) = (1 + kd\ \ 1\ \ 1 + md) \in G(S)$. It follows that $(1\ \ 1 + kd\ \ 1 + md) = (1 + kd\ \ 1\ \ 1 + md)^2 \in G(S)$.

Finally, given distinct nonnegative integers k, m, n, without loss of generality, we can assume that $0 < k < m < n$. Then we have

$$(1\ \ 1+(m-1)d\ \ 1+(n-1)d) \in G(S)$$

https://doi.org/10.4153/S0008439519000250 Published online by Cambridge University Press

by the second claim, and so $(1+d\ \ 1+md\ \ 1+nd) \in G(S)$. Thus we suppose $k > 1$. Then we have that $\tau = (1\ \ 1+d\ \ 1+nd) \in G(S)$ by the first claim and $(1\ \ 1+kd\ \ 1+md) \in G(S)$ by the second claim. Hence

$$\begin{aligned}\tau^{-1}(1\ \ 1+kd\ \ 1+md)\tau &= (\tau^{-1}(1)\ \ \tau^{-1}(1+kd)\ \ \tau^{-1}(1+md)) \\ &= (1+nd\ \ 1+kd\ \ 1+md).\end{aligned}$$

Thus $\rho = (1+kd\ \ 1+md\ \ 1+nd) \in G(S)$. ■

Denote by $\mathrm{Alt}(\mathbb{N})$ the alternating group on $\mathbb{N}$. Then $\mathrm{Alt}(\mathbb{N})$ has index 2 in the group $\mathrm{FSym}(\mathbb{N})$ and it is generated by all 3-cycles on $\mathbb{N}$ [6, Exercise 1.6.8].

Corollary 2.8 *Let S be a permutable set. Then the following are equivalent.*

(i) $d(S) = 1$.
(ii) $(1\ 2\ 3) \in G(S)$.
(iii) $\mathrm{Alt}(\mathbb{N}) \subset G(S)$.

Proof Clearly, (iii) ⇒ (ii) ⇒ (i). If $d(S) = 1$, then Lemma 2.7 implies that $G(S)$ contains all 3-cycles, and so $\mathrm{Alt}(\mathbb{N}) \subset G(S)$. This proves (i) ⇒ (iii). ■

Lemma 2.9 *Let $\sigma = (1\ \ 1+km\ \ 1+2km)$ with $k, m \in \mathbb{N}$. Then $(1\ \ 1+m\ \ 1+2m) \in G(S^k)$ if and only if $\sigma_{[0]}\sigma_{[1]}\cdots\sigma_{[k-1]} \in G(S)$.*

Proof The proof is straightforward. ■

Lemma 2.10 *If $d(S) = d$, then $d(S^k) = d/\gcd(d,k)$ for any $k \in \mathbb{N}$.*

Proof Let $\gcd(d,k) = d_1$, $d = pd_1$, $k = qd_1$, and $d(S^k) = t$. We only need to prove that $t = p$.

For $\sigma = (1\ \ 1+kp\ \ 1+2kp)$, we have $\sigma = (1\ \ 1+dq\ \ 1+2dq) \in G(S)$ by Lemma 2.7, and so $\sigma_{[0]}\sigma_{[1]}\cdots\sigma_{[k-1]} \in G(S)$. Thus $(1\ \ 1+p\ \ 1+2p) \in G(S^k)$ by Lemma 2.9. Hence $t \mid p$ by Lemma 2.4. For $\sigma = (1\ \ 1+kt\ \ 1+2kt)$, we have $\tau = \sigma_{[0]}\sigma_{[1]}\cdots\sigma_{[k-1]} \in G(S)$ by Lemma 2.9, since $(1\ \ 1+t\ \ 1+2t) \in G(S^k)$ by Lemma 2.7. Noting that $\tau(1) - 1 = kt = qd_1t$, we have $d \mid qd_1t$ by Lemma 2.4. Since $\gcd(p,q) = 1$, we have $p \mid t$. This finishes the proof. ■

A subset S of a semigroup is called a nil set of bounded index if there exists a positive integer s such that $a^s = 0$ holds for all $a \in S$.

Theorem 2.11 *Let $d(S) = d$ and $k \in \mathbb{N}$.*

(i) *If $d \mid k$, then $d(S^k) = 1$.*
(ii) *If $\gcd(d,k) = 1$ and S is a nil set of bounded index, then elements in S^k are nilpotent.*

Proof (i) follows immediately from Lemma 2.10.

(ii) Suppose S is a nil set of bounded index, that is, $a^s = 0$ for all $a \in S$ and for some $s \in \mathbb{N}$. If k is prime to d, then $kv = du + 1$ for some $u, v \in \mathbb{N}$. Let $d' = du$ and $k' = kv$. Then $k' = d' + 1$, and moreover, we can assume that $d' \geq s$.

Let $N_i = \{i + jd \mid j \in \{0\} \cup \mathbb{N}\}$ and let G_i be the subgroup generated by

$$\{(i{+}jd \;\; i{+}md \;\; i{+}nd) \mid j, m, n \text{ are distinct nonnegative integers}\}$$

for $i = 1, 2, \ldots, d$. Then $G_i \subset G(S)$ by Lemma 2.7. Furthermore, G_i acts transitively on N_i and $\operatorname{supp}(\sigma) \subset N_i$ for all $\sigma \in G_i$. Thus there exists $\sigma_i \in G_i$ such that $\sigma_i(i) = i + (i-1)d'$ for $i = 1, 2, \ldots, d'$, and the supports of $\sigma_1, \sigma_2, \ldots, \sigma_{d'}$ are disjoint. Let $\sigma = \sigma_1\sigma_2\cdots\sigma_{d'} \in G(S)$ and let i_m be the maximum in $\operatorname{supp}(\sigma)$. Take $r \in \mathbb{N}$ such that $r > \max\{k'd', i_m\}$. Then for all sufficiently large integers t and for all $u, v \in S^t$, $a_1, a_2, \ldots, a_r \in S$ with $a_{qk'+i} = a_i$ for integers $q \geq 0$ and $1 \leq i \leq k'$ we have

$$\begin{aligned} u(a_1\cdots a_{k'})^{d'}\cdots a_r v &= ua_1a_2\cdots a_{k'd'}\cdots a_r v \\ &= ua_{\sigma(1)}a_{\sigma(2)}\cdots a_{\sigma(k'd')}\cdots a_{\sigma(r)}v \\ &= ua_1a_{2+d'}\cdots a_{d'+(d'-1)d'}a_{\sigma(d'+1)}\cdots a_{\sigma(k'd')}\cdots a_{\sigma(r)}v \\ &= ua_1^{d'}a_{\sigma(d'+1)}\cdots a_{\sigma(k'd')}\cdots a_{\sigma(r)}v \\ &= 0, \end{aligned}$$

since $a_{(i-1)d'+i} = a_{(i-1)k'+1} = a_1$ for $i = 1, 2, \ldots, d'$. Thus elements in $S^{k'}$ are nilpotent, and so elements in S^k are nilpotent. ■

Remark 2.12 The subsemigroup generated by a permutable set is not necessarily permutable. For example, let $S = \{e_{12}, e_{21}\}$. Then S is $(1\,3)$-permutable for $(1\,3) \in S_3$ and the subsemigroup $T = \{0, e_{11}, e_{22}, e_{12}, e_{21}\}$ generated by S is not permutable. Suppose otherwise. Then $(1\,2) \in G(T)$ by [12, Theorem 1]. It follows that $ue_{12}e_{22}v = ue_{22}e_{12}v = 0$ for all sufficiently large integers t and for all $u, v \in S^t$. Particularly, $e_{12} = e_{11}^t e_{12}e_{22}e_{22}^t = 0$, a contradiction.

It is interesting to compare Theorem 2.11(i) with [12, Theorem 1], which states that $(1\,2) \in G(S)$ or equivalently $G(S) = \mathrm{FSym}(\mathbb{N})$ if S is a permutable semigroup. Now we can deduce [12, Theorem 1] from our results.

Corollary 2.13 *Suppose S is a permutable semigroup. Then $G(S) = \mathrm{FSym}(\mathbb{N})$.*

Proof Suppose $d(S) = d$. Then $(1\;2\;3) \in G(S^d)$ by Theorem 2.11(i) and Corollary 2.8. Since S is a semigroup, we have $S^{d+1} \subset S^d$. Thus

$$\begin{aligned} &u(a_1\cdots a_d)(a_{d+1}\cdots a_{2d+1})(a_{2d+2}\cdots a_{3d+1})v \\ &\qquad = u(a_{d+1}\cdots a_{2d+1})(a_{2d+2}\cdots a_{3d+1})(a_1\cdots a_d)v \end{aligned} \tag{2.1}$$

for all sufficiently large $t \in \mathbb{N}$ and for all $u, v \in S^t$ and $a_1, a_2, \ldots, a_{3d+1} \in S$. Let σ be the permutation appearing in equation (2.1). Then $\sigma(1) - 1 = d$ and $\sigma(3d+1) - (3d+1) = -2d-1$, whence $d(S) = 1$ by Lemma 2.4. In the case of $d = 1$, equation (2.1) gives $\sigma = (1\,2\,3\,4)$, which is odd. Thus $G(S)$ contains an odd permutation and, by Corollary 2.8, the even permutations. Hence $G(S) = \mathrm{FSym}(\mathbb{N})$. ■

3 Simultaneous Triangularization of Permutable Matrices

It is well known that an associative algebra R over a field K is a Lie algebra over K with the Lie product defined by the commutator $[a, b] = ab - ba$.

Lemma 3.1 *Let R be an associative algebra over a field with a generating set S. If S is σ-permutable for $\sigma = (1\,2\,3) \in S_3$, then R is a nilpotent Lie algebra.*

Proof It suffices to prove that

$$[[a_1 \cdots a_r, b_1 \cdots b_s], c_1 \cdots c_t] = 0 \tag{3.1}$$

for all $r, s, t \in \mathbb{N}$ and $a_1, \ldots, a_r, b_1, \ldots, b_s, c_1, \ldots, c_t \in S$. Since S is σ-permutable, we have $abc = bca$ and so $[a, bc] = 0$ and $[a, [b, c]] = 0$ for all $a, b, c \in S$. Therefore, S^{2k} is contained in the center of R for all $k \in \mathbb{N}$. If one of r, s, t is even, then (3.1) holds. If r, s, t are all odd, then

$$[[a_1 \cdots a_r, b_1 \cdots b_s], c_1 \cdots c_t] = a_1 \cdots a_{r-1} b_1 \cdots b_{s-1} c_1 \cdots c_{t-1} [[a_r, b_s], c_t] = 0,$$

as desired. ∎

Lemma 3.2 *Let R be a finite-dimensional associative algebra over a field with a generating set S. If S is permutable and $d(S) = 1$, then R/J is commutative, where J is the Jacobson radical of R.*

Proof If $d(S) = 1$, then $(1\,2\,3) \in G(S)$ by Lemma 2.6. It follows that $uabcv = ubcav$ for all sufficiently large integers t and for all $u, v \in S^t$ and $a, b, c \in S$, which implies that $abc - bca \in J$. Let $\bar{R} = R/J$ and $\bar{S}$ be the image of S under the natural homomorphism $R \to \bar{R}$. Then $\bar{R}$ is an algebra generated by $\bar{S}$ and $\bar{S}$ is σ-permutable for $\sigma = (1\,2\,3) \in S_3$. By Lemma 3.1, $\bar{R}$ is Lie nilpotent, and so $\bar{R}$ is commutative. ∎

Lemma 3.3 ([14, Theorem 1.3.2]) *An algebra of matrices is triangularizable if and only if each commutator $BC - CB$ is nilpotent.*

Theorem 3.4 *Let S be a set of $n \times n$ matrices over an algebraically closed field. If S is permutable and $d(S) = 1$, then S is triangularizable.*

Proof Let R be the subalgebra generated by S. By Lemma 3.2 and Lemma 3.3, R is triangularizable, and so is S. ∎

For a nonidentity permutation $\sigma \in S_n$, define

$$\Delta(\sigma) = \gcd\left(|\sigma(1) - 1|, |\sigma(2) - 2|, \ldots, |\sigma(n) - n| \right).$$

Theorem 3.5 *Let S be a set of $n \times n$ matrices over an algebraically closed field. If S is σ-permutable for some permutation σ with $\Delta(\sigma) = 1$, then S is triangularizable.*

Proof This follows directly from Lemma 2.4 and Theorem 3.4. ∎

Remark 3.6 The condition $\Delta(\sigma) = 1$ in Theorem 3.5 is not necessary. For example, the set S of invertible upper triangular 2×2 matrices over a field is triangularizable but not permutable, since, if S is permutable, it must be commutative by [12, Theorem 1].

Using Theorem 3.5, we can now generalize Radjavi's result [13, Theorem 1]. The proof is based on ideas in the proof of [13, Theorem 1].

Theorem 3.7 *Let S be a set of $n \times n$ matrices over a field K of characteristic either zero or greater than $n/2$. Suppose that the eigenvalues of every matrix in S belong to K. If there exists a permutation $\sigma \in S_r$ with $\Delta(\sigma) = 1$ such that*

$$\operatorname{tr}\big((A_1\cdots A_r - A_{\sigma(1)}\cdots A_{\sigma(r)})B_1\cdots B_s\big) = 0 \tag{3.2}$$

for all $s \in \mathbb{N}$ and for all $A_1,\dots,A_r,B_1,\dots,B_s \in S$, then S is triangularizable.

Proof Without loss of generality we can assume that the field K is algebraically closed and that

$$\operatorname{tr}(A_1\cdots A_r - A_{\sigma(1)}\cdots A_{\sigma(r)}) = 0, \tag{3.3}$$

since otherwise we can replace r with $r+1$ and σ with $\sigma_{[1]}$. For, clearly $\Delta(\sigma_{[1]}) = 1$. Using (3.2) and the fact that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ for all $n\times n$ matrices A, B over K, we have

$$\operatorname{tr}\big((B_sA_1\cdots A_r - B_sA_{\sigma(1)}\cdots A_{\sigma(r)})B_1\cdots B_{s-1}\big) = 0, \tag{3.4}$$

for all $s \in \mathbb{N}$ and for all $A_1,\dots,A_r,B_1,\dots,B_s \in S$. Now rewriting A_i as A_{i+1}, for $i = 1,2,\dots,r$ and B_s as A_1 in (3.4), we have

$$\operatorname{tr}\big((A_1A_2\cdots A_{r+1} - A_{\tau(1)}A_{\tau(2)}\cdots A_{\tau(r+1)})B_1\cdots B_{s-1}\big) = 0, \quad (\tau = \sigma_{[1]}),$$

for all $s \in \mathbb{N}$ and for all $A_1,\dots,A_{r+1},B_1,\dots,B_{s-1} \in S$, implying that (3.2) and (3.3) hold with r replaced by $r+1$ and σ by $\sigma_{[1]}$.

Let R be the unital subalgebra generated by S over K. Then

$$\operatorname{tr}\big((A_1\cdots A_r - A_{\sigma(1)}\cdots A_{\sigma(r)})C\big) = 0,$$

for all $A_1,\dots,A_r \in S$ and $C \in R$. By [14, Theorem 1.5.1] or [20, Theorem 1], we can assume that R consists of block upper triangular matrices with k diagonal blocks. We must prove all diagonal blocks are 1×1 matrices. Suppose that, on the contrary, the first diagonal block, say, is of order $m > 1$. Let $C^{(i)}$ denote the i-th diagonal block of $C \in R$, R_i the algebra consisting of all the matrices $C^{(i)}$, and S_i the set of i-th diagonal blocks of members in S. By [14, Theorem 1.5.1], there exists a subset J of $\{1,2,\dots,k\}$ containing 1 such that

(i) R_1 is the algebra of the $m\times m$ matrices over K;
(ii) $A^{(i)} = A^{(j)}$ for all A in R and $i, j \in J$;
(iii) for any $C \in R_1$, there is an $A \in R$ such that $A^{(i)} = C$ for $i \in J$ and $A^{(i)} = 0$ for $i \notin J$, $i = 1,2,\dots,k$.

Denote $A^{(1)}$ briefly by A' for all $A \in R$. Then we have that

$$t\operatorname{tr}\big((A'_1\cdots A'_r - A'_{\sigma(1)}\cdots A'_{\sigma(r)})C\big) = \operatorname{tr}\big((A_1\cdots A_r - A_{\sigma(1)}\cdots A_{\sigma(r)})A\big) = 0,$$

https://doi.org/10.4153/S0008439519000250 Published online by Cambridge University Press

where t is the cardinality of the set J, $A_1, \dots, A_r \in S$, and $C \in R_1$ and $A \in R$ are as in (iii) above. By (ii) above, $mt \le n$, and so $t \le n/m \le n/2$. Thus $\operatorname{tr}\big((A'_1 \cdots A'_r - A'_{\sigma(1)} \cdots A'_{\sigma(r)})C\big) = 0$ for all $A_1, \dots, A_r \in S$ and $C \in R_1$, since the characteristic of K is 0 or greater than $n/2$. Thus

$$A'_1 \cdots A'_r - A'_{\sigma(1)} \cdots A'_{\sigma(r)} = 0,$$

for all $A_1, \dots, A_r \in S$, since R_1 is the algebra of the $m \times m$ matrices by (i) above. Note that S_1 is a generating set of R_1. Then Theorem 3.5 implies that R_1 is triangularizable, and so $m = 1$, contradicting the assumption that $m > 1$. ■

4 Simultaneous Triangularization of Permutable Subspace of Nilpotent Matrices

It is well known that a set S of nilpotent matrices is triangularizable if and only if $S^k = 0$ for some $k \in \mathbb{N}$ [14, Theorem 2.1.7].

Theorem 4.1 *Let S be a subset of nilpotent $n \times n$ matrices over a field K. Then S is triangularizable if and only if S is σ-permutable for some permutation σ with $\Delta(\sigma) = 1$.*

Proof *Sufficiency*. Let $\bar{K}$ denote the algebraic closure of K. Then S is triangularizable over $\bar{K}$ by Theorem 3.5, and so $S^k = 0$ for some $k \in \mathbb{N}$. It follows that S is triangularizable over K.

Necessity. Suppose S is triangularizable. Then $S^k = 0$ for some $k \in \mathbb{N}$, which implies that S is σ-permutable for all nonidentical $\sigma \in S_k$. ■

The following example shows that for every integer $d > 1$ there exists a $(1\ d{+}1)$-permutable set S of nilpotent matrices such that S is not triangularizable.

Example 4.2 *Let $S = \{e_{12}, e_{23}, \dots, e_{d1}\}$ for $d > 1$. Then S is a σ-permutable set of square-zero matrices for $\sigma = (1\ d{+}1) \in S_{d+1}$, but S is not triangularizable.*

Proof Indeed, let $a_k = e_{i_k j_k}$, where $1 \le i_k, j_k \le d$ and $j_k \equiv i_k + 1 \mod d$ for $k = 1, 2, \dots, d+1$. We first claim that if $a_1 a_2 \cdots a_{d+1} \ne 0$, then $a_1 = a_{d+1}$. Indeed, in this case we must have

$$i_k + 1 \equiv i_{k+1} \mod d, \ k = 1, 2, \dots, d.$$

It follows that $i_1 \equiv i_{d+1} \mod d$, which implies that $i_1 = i_{d+1}$ since $1 \le i_1, i_{d+1} \le d$. Thus $a_1 = a_{d+1}$.

Let $\sigma = (1\ d{+}1) \in S_{d+1}$. If $a_1 \ne a_{d+1}$, then

$$a_1 a_2 \cdots a_{d+1} = a_{\sigma(1)} a_{\sigma(2)} \cdots a_{\sigma(d+1)}$$

since both sides of the last equation equal 0 by the claim above. If $a_1 = a_{d+1}$, then $a_1 a_2 \cdots a_{d+1} = a_{\sigma(1)} a_{\sigma(2)} \cdots a_{\sigma(d+1)}$. Thus S is σ-permutable.

Since $e_{12} e_{23} \cdots e_{d1} = e_{11}$, we see that S is not triangularizable. ■

Lemma 4.3 *Let S be a subset of an algebra over a field K. If S is σ-permutable for some $\sigma \in S_n$, the subspace spanned by S is σ-permutable.*

Proof Given $a_1, \dots, a_m \in S$ and $\lambda_{kj} \in K$ for $k = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$, let $u_k = \sum_{j=1}^{m} \lambda_{kj} a_j$ for $k = 1, 2, \dots, n$. Then

$$\begin{aligned} u_1 \cdots u_n &= \sum \lambda_{1j_1} \cdots \lambda_{nj_n} a_{j_1} \cdots a_{j_n} \\ &= \sum \lambda_{\sigma(1)j_{\sigma(1)}} \cdots \lambda_{\sigma(n)j_{\sigma(n)}} a_{j_{\sigma(1)}} \cdots a_{j_{\sigma(n)}} \\ &= u_{\sigma(1)} \cdots u_{\sigma(n)} \end{aligned}$$

Thus the subspace spanned by S is σ-permutable. ∎

Lemma 4.4 *Let A and B be $n \times n$ matrices over a field K that are simultaneously triangularizable over the algebraic closure of K. If AB and $A + B$ are nilpotent, then A and B are nilpotent.*

Proof Without loss of generality, we assume that A and B are upper triangular matrices. Let a_i and b_i be the i-th diagonal entries of A and B, respectively. If AB and $A + B$ are nilpotent, then $a_i b_i = 0$ and $a_i + b_i = 0$, which force $a_i = b_i = 0$. Thus A and B are strictly upper triangular matrices, and so they are nilpotent. ∎

Theorem 4.5 *Let S be a subspace of nilpotent $n \times n$ matrices over a field K. If S is permutable and K contains at least $d(S)$ elements, then S is triangularizable.*

Proof If $d(S) = 1$, it follows immediately from Theorem 4.1. Suppose on the contrary that $d(S) = d > 1$. Denote by U the subspace spanned by S^d. Then U is permutable and $d(U) = 1$ by Lemma 4.3 and Theorem 2.11. It follows from Theorem 3.4 that U is triangularizable over the algebraic closure of K. Let U_0 be the set of nilpotent matrices in U. Then U_0 is a K-subspace of U such that $A^d \in U_0$ for all $A \in S$, and CD and DC are nilpotent for all $C \in U$ and $D \in U_0$.

Let $T_m = \{A \in S \mid A^m S^{d-m} \subset U_0\}$ for $m = 1, 2, \dots, d$. We now claim that

$$S = T_d = \cdots = T_2 = T_1.$$

We proceed via a reverse inductive argument. Since $A^d \in U_0$, we have that $T_d = S$. Suppose $S = T_m$ for $1 < m \le d$. Clearly, $S = T_d \supset \cdots T_2 \supset T_1$. To prove $S = T_{m-1}$, we need to prove that $A^{m-1}BW \in U_0$ for all $A, B \in S$ and $W \in S^{d-m}$. Since S is a subspace, we have $(A + zB)^m W \in U_0$. Expanding $(A + zB)^m W$ we get

$$A^m W + zMW + \cdots + z^{m-1}NW + z^m B^m W \equiv 0 \mod U_0,$$

for all $z \in K$, where $M = \sum_{i=0}^{m-1} A^i BA^{m-i-1}$ and $N = \sum_{i=0}^{m-1} B^i AB^{m-i-1}$. Since $A^m W \equiv B^m W \equiv 0 \mod U_0$ by the inductive hypothesis, we have

$$zMW + \cdots + z^{m-1}NW \equiv 0 \mod U_0. \tag{4.1}$$

Since K contains at least d elements and $m \le d$, we can choose distinct nonzero elements $z_i \in K$ $(1 \le i \le m - 1)$ and replace z with z_i in (4.1) to obtain

$$z_i MW + \cdots + z_i^{m-1} NW \equiv 0 \mod U_0, \quad i = 1, 2, \dots, m - 1.$$

Now we get $MW \equiv 0 \mod U_0$ by a Vandermonde determinant argument. Write

$$MW = CAW + A^{m-1}BW, \tag{4.2}$$

where $C = \sum_{i=0}^{m-2} A^i BA^{m-i-2}$. Note that $AWA^{m-1} \in U_0$ and $BWC \in U$. Then $AWA^{m-1}BWC$ is nilpotent, and so $CAWA^{m-1}BW$ is nilpotent. Applying Lemma 4.4 to (4.2), we get CAW and $A^{m-1}BW$ are nilpotent. Hence $A^{m-1}BW \in U_0$, as desired.

Now $S = T_1$ implies that $S^d \subset U_0$. Thus $S^{dn} = 0$ and so $d(S) = 1$, which contradicts the assumption $d(S) > 1$. ∎

Remark 4.6 Theorem 4.5 fails for the field K of two elements if the condition that K contains at least $d(S)$ elements is removed. Let

$$S = \left\{ ae_{12} + be_{23} + (a+b)e_{31} \mid a, b \in K \right\}.$$

Then S is a subspace. Noting that products of any three matrices in S are diagonal, we can see that S consists of nilpotent matrices, S is permutable, and S is not triangularizable.

5 Linear Triangularization of Polynomial Maps

Let K be a field of characteristic zero and $K[X] = K[x_1, x_2, \dots, x_n]$ be the polynomial algebra in the variables $x_1, x_2, \dots, x_n$ over K.

A polynomial map is an n-tuple $F = (F_1, F_2, \dots, F_n) \in K[X]^n$ and its Jacobian matrix is $JF = (\partial F_i / \partial x_j)$. A polynomial map F is called linearly triangularizable if there exists an invertible linear map $L \in \mathrm{GL}_n(K)$ such that $L^{-1}FL$ is upper triangular, in the sense that $F_i - x_i \in k[x_{i+1}, \dots, x_n]$ for all $1 \le i \le n-1$ and $F_n = x_n$.

For $F \in K[X]^n$, write $JF(K)$ for the set of Jacobian matrices $JF|_{X=u}$, $u \in K^n$.

Lemma 5.1 ([19, 22]) *A polynomial map $F = X + H$ with JH nilpotent is linearly triangularizable if and only if the set $JH(K)$ is simultaneously triangularizable.*

By the lemma above it is natural to study linear triangularization of polynomial maps from the point of view of simultaneous triangularization of matrices.

Yan and Tang [21, Theorem 1] proved that a polynomial map $F = X + H$ with JH nilpotent is linearly triangularizable if the set $JH(K)$ is σ-permutable for all $\sigma \in S_r$ for some $r \ge 2$. We now give generalizations of Theorem 2.1 and Corollary 2.3 in [21] that are direct corollaries of Theorem 4.5 and Lemma 5.1.

Theorem 5.2 *A polynomial map $F = X + H$ with JH nilpotent is linearly triangularizable if and only if the set $JH(K)$ is σ-permutable for some $\sigma \in S_r$ with $\Delta(\sigma) = 1$.*

Theorem 5.3 *Let $F = X + H$ be a polynomial map such that JH is nilpotent. Write $JH = \sum_\alpha A_\alpha x_\alpha$ and let S be the set of the A_α's. Then F is linearly triangularizable if and only if S is σ-permutable for some $\sigma \in S_r$ with $\Delta(\sigma) = 1$.*

A Jacobian matrix JH is called additive-nilpotent if each matrix in the subspace spanned by $JH(K)$ is nilpotent [7, 16].

The following result is a direct corollary of Theorem 4.5 and Lemma 5.1.

Theorem 5.4 *A polynomial map $F = X + H$ with JH additive-nilpotent is linearly triangularizable if and only if the set $JH(K)$ is permutable.*

https://doi.org/10.4153/S0008439519000250 Published online by Cambridge University Press

Acknowledgements The authors would like to thank the referees for helpful comments.

References

[1] H. Bass, E. H. Connell, and D. Wright, *The Jacobian Conjecture: Reduction of degree and formal expansion of the inverse*. Bull. Amer. Math. Soc. 7(1982), no. 2, 287–330. https://doi.org/10.1090/S0273-0979-1982-15032-7
[2] C. C. Cheng, *Cubic linear Keller maps*. J. Pure Appl. Algebra 160(2001), no. 1, 13–19. https://doi.org/10.1016/S0022-4049(00)00076-1
[3] M. de Bondt, *The strong nilpotency index of a matrix*. Linear Multilinear Algebra 62(2014), no. 4, 486–497. https://doi.org/10.1080/03081087.2013.784282
[4] M. de Bondt and A. van den Essen, *The Jacobian Conjecture: Linear triangularization for homogeneous polynomial maps in dimension three*. J. Algebra 294(2005), no. 1, 294–306. https://doi.org/10.1016/j.jalgebra.2005.04.018
[5] M. A. Dehghan and M. Radjabalipour, *Matrix algebras and Radjavi's trace condition*. Linear Algebra Appl. 148(1991), 19–25. https://doi.org/10.1016/0024-3795(91)90083-9
[6] J. D. Dixon and B Mortimer, *Permutation groups*, Graduate Texts in Mathematics, 163, Springer-Verlag, New York, 1996. https://doi.org/10.1007/978-1-4612-0731-3
[7] H. Guo, M. de Bondt, X. Du, and X. Sun, *Polynomial maps with invertible sums of Jacobian matrices and directional derivatives*. Indag. Math. 23(2012), no. 3, 256–268. https://doi.org/10.1016/j.indag.2011.11.007
[8] D. Hadwin, *Radjavi's trace condition for triangularizability*. J. Algebra 109(1987), no. 1, 184–192. https://doi.org/10.1016/0021-8693(87)90172-4
[9] D. Liu, X. Du, and X. Sun, *Quadratic linear Keller maps of nilpotency index three*. Linear Algebra Appl. 429(2008), no. 1, 12–17. https://doi.org/10.1016/j.laa.2008.01.032
[10] G. H. Meisters and C. Olech, *Strong nilpotence holds in dimensions up to five only*. Linear Multilinear Algebra 30(1991), no. 2, 231–255. https://doi.org/10.1080/03081089108818109
[11] K. Pate and C. C. Cheng, *Quadratic homogeneous Keller maps of rank two*. Linear Algebra Appl. 476(2015), 16–27. https://doi.org/10.1016/j.laa.2015.02.014
[12] M. S. Putcha and A. Yaqub, *Semigroups satisfying permutation identities*. Semigroup Forum 3(1971), no. 1, 68–73. https://doi.org/10.1007/BF02572944
[13] H. Radjavi, *A trace condition equivalent to simultaneous triangularizability*. Canad. J. Math. 38(1986), no. 2, 376–386. https://doi.org/10.4153/CJM-1986-018-1
[14] H. Radjavi and P. Rosenthal, *Simultaneous triangularization*, Springer-Verlag, New York, 2000. https://doi.org/10.1007/978-1-4612-1200-3
[15] K. Rusek, *Linearly triangularizable quadratic endomorphisms*. Linear Multilinear Algebra 64(2016), no. 8, 1527–1537. https://doi.org/10.1080/03081087.2015.1099601
[16] X. Sun, *On additive-nilpotency of Jacobian matrices of polynomial maps*. Linear Algebra Appl. 439(2013), no. 12, 3746–3751. https://doi.org/10.1016/j.laa.2013.10.014
[17] X. Sun, *On the strong nilpotence problem*. Algebra Colloq. 21(2014), no. 1, 117–128. https://doi.org/10.1142/S100538671400008X
[18] A. van den Essen, *Polynomial automorphisms and the Jacobian conjecture*, Progress in Mathematics, 190, Birkhäuser, Berlin, 2000. https://doi.org/10.1007/978-3-0348-8440-2
[19] A. van den Essen and E. Hubbers, *Polynomial maps with strongly nilpotent Jacobian matrix and the Jacobian conjecture*. Linear Algebra Appl. 247(1996), no. 6, 121–132. https://doi.org/10.1016/0024-3795(95)00095-X
[20] J. F. Watters, *Block triangularization of algebras of matrices*. Linear Algebra Appl. 32(1980), 3–7. https://doi.org/10.1016/0024-3795(80)90003-8
[21] D. Yan and G. Tang, *The linear triangularizability of some Keller maps*. Linear Algebra Appl. 438(2013), no. 9, 3649–3653. https://doi.org/10.1016/j.laa.2012.12.018
[22] Jie-Tai Yu, H. Bass, E. H. Connell, and D. Wright, *On generalized strongly nilpotent matrices*. Linear Multilinear Algebra 41(1996), no. 1, 19–22. https://doi.org/10.1080/03081089608818457

College of Mathematics and Information Science, Henan Normal University, Xinxiang 453007, China
e-mail: a892691755@139.com

School of Mathematics, Liaoning Normal University, Dalian 116029, China
e-mail: tiantian8835@163.com

School of Mathematics, Jilin University, Changchun 130012, China
e-mail: duxk@jlu.edu.cn