On the Mesyan conjecture

Dedicated to Professor Vesselin Drensky on the occasion of his 70th birthday.

Pedro Souza FAGUNDES[1],[*] , **Thiago Castilho de MELLO**[2] ,
Pedro Henrique da Silva dos SANTOS[2]
[1]Institute of Mathematics, Statistics and Scientific Computing, State University of Campinas,
Campinas, SP, Brazil
[2]Instituto de Ciência e Tecnologia, Universidade Federal de São Paulo, São José dos Campos, SP, Brazil

Abstract: The well-known Lvov–Kaplansky conjecture states that the image of a multilinear polynomial f evaluated on $n \times n$ matrices is a vector space. A weaker version of this conjecture, known as the Mesyan conjecture, states that if $m = \deg f$ and $n \geq m - 1$ then its image contains the set of trace zero matrices. Such conjecture has been proved for polynomials of degree $m \leq 4$. The proof of the case $m = 4$ contains an error in one of the lemmas. In this paper, we correct the proof of such lemma and present some evidence which allows us to state the Mesyan conjecture for the new bound $n \geq \frac{m+1}{2}$, which cannot be improved.

Key words: Images of polynomials, Lvov–Kaplansky conjecture, Mesyan conjecture, polynomial identities, central polynomials

1. Introduction

Let K be a field and let $M_n(K)$ denote the algebra of $n \times n$ matrices over K. A famous problem known as Lvov–Kaplansky conjecture asserts: the image of a multilinear polynomial (in noncommutative variables) on $M_n(K)$ is a vector space. Such conjecture is equivalent to the following: the image of a multilinear polynomial on $M_n(K)$ is $\{0\}$, K (viewed as the set of scalar matrices), $sl_n(K)$ (the set of traceless matrices) or $M_n(K)$.

Although proving that some subset is a vector space seems to be at first look a simple problem, a solution to Lvov–Kaplansky conjecture is known only for $n = 2$ [11, 15]. The case $n = 3$ has interesting progress, but not a solution [12]. This conjecture motivated other studies related to images of polynomials. For instance, papers on images of polynomials on some subalgebras of $M_n(K)$, images of Lie, and Jordan polynomials on Lie and Jordan algebras have been published since then (see [3, 7–9, 13, 16, 17]). For a nice compilation of results on images of polynomials, we recommend the survey [14].

An analogue of the Lvov–Kaplansky conjecture for the infinite dimensional case, i.e., for the algebra $A = \text{End}(V)$, where V is a countably infinite-dimensional vector space over K, was studied in [21]. In that paper, the author proved that if f is any nonzero multilinear polynomial, than the image of f is A.

A weakening of the Lvov–Kaplansky conjecture is the so called Mesyan conjecture [18, Conjecture 11]:

---

[*]Correspondence: pedro.fagundes@ime.unicamp.br

Conjecture 1.1 *Let K be a field, $n \geq 2$ and $m \geq 1$ be integers and $f(x_1, \ldots, x_m)$ a nonzero multilinear polynomial in $K\langle x_1, \ldots, x_m \rangle$. If $n \geq m - 1$, then the image of f on $M_n(K)$ contains all trace zero matrices.*

The above conjecture is based on the following result (see [18, Proposition 10]).

Proposition 1.2 *Let K be a field, $n \geq 2$ and $m \geq 1$ be integers, and $f(x_1, \ldots, x_m)$ be a nonzero multilinear polynomial in $K\langle x_1, \ldots, x_m \rangle$. If $n \geq m - 1$, then the K-subspace of $M_n(K)$ generated by the image of f contains $sl_n(K)$.*

In fact, once one assumes the Lvov–Kaplansky conjecture is true, the sentence "the image of f on $M_n(K)$ contains $sl_n(K)$" is equivalent to "f is not an identity nor a central polynomial of $M_n(K)$". On the other hand, it is well-known that $M_n(K)$ has no identities or central polynomials of degree $m \leq n + 1$, and this makes Conjecture 1.1 a particular case of the Lvov–Kaplansky conjecture. In particular, a counter-example to Mesyan conjecture is a counter-example for the Lvov-Kaplanksy conjecture.

In this paper, we present a more general result than Proposition 1.2, and restate the conjecture for a more general case (see Conjecture 3.3 below). We also discuss the relationship of Conjecture 3.3 to minimal degrees of central polynomials and identities for $M_n(K)$.

Positive solutions to Conjecture 1.1 have been presented for the algebra $M_\infty(K)$, of finitary matrices and for $m \leq 4$. In [20] the author proved an analogue of the Mesyan conjecture for algebra $M_\infty(K)$, namely, if K is an infinite field and f is a nonzero multilinear polynomial, then the image of f on $M_\infty(K)$ contains $sl_\infty(K)$ (the set of trace zero finitary matrices).

The case $m = 2$ is a direct consequence of results of Shoda [19] (for the characteristic zero case) and Albert and Muckenhoupt [1] (for the positive characteristic case), where they prove that any trace zero matrix is given by a commutator of two matrices, while the case $m = 3$ was proved by Mesyan himself in [18].

The case $m = 4$ was presented by Buzinsky and Winstanley in [5], but their proof contains a crucial error in one of its lemmas, so the solution is not correct. In this paper, we present a correction for that lemma, confirming the Mesyan conjecture for $m = 4$.

2. Preliminaries

In this section we define the basic objects and present the basic results necessary to the paper. We start with the definition of a multilinear polynomial. Throughout the paper, unless otherwise stated, K will denote an arbitrary field and all algebras are considered over K.

Definition 2.1 *Let m be a positive integer. By $K\langle x_1, \ldots, x_m \rangle$ we denote the free associative K-algebra, freely generated by $\{x_1, \ldots, x_m\}$. The elements of $K\langle x_1, \ldots, x_m \rangle$ will be called* polynomials *in the noncommutative variables $x_1, \ldots, x_m$. A polynomial $f(x_1, \ldots, x_m)$ is said to be* multilinear *if it can be written as*

$$f(x_1, \ldots, x_m) = \sum_{\sigma \in S_m} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)},$$

where S_m denotes the group of permutations of $\{1, \ldots, m\}$ and $\alpha_\sigma \in K$, for $\sigma \in S_m$.

For a given K-algebra A, a polynomial $f(x_1, \ldots, x_m) \in K\langle x_1, \ldots, x_m\rangle$ defines a map (also denoted by f)

$$
\begin{array}{cccc}
f: & A^m & \longrightarrow & A \\
& (a_1, \ldots, a_m) & \mapsto & f(a_1, \ldots, a_m)
\end{array}
$$

The image of such map f is called *the image of the polynomial f on A*, and will be denoted by $f(A)$.

Some well-known properties of the set $f(A)$ are given below.

Proposition 2.2 *Let A be an algebra and $f(x_1, \ldots, x_m) = \sum_{\sigma \in S_n} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)}$ be a multilinear polynomial, where $\alpha_\sigma \in K$, for all $\sigma \in S_n$. Then*

1. *$f(A)$ is closed under automorphisms of A. In particular, $f(A)$ is closed under conjugation by invertible elements.*

2. *$f(A)$ is closed under scalar multiplication.*

3. *The linear span of $f(A)$ is a Lie ideal of A.*

4. *If $\sum_{\sigma \in S_m} \alpha_\sigma \neq 0$ then $f(A) = A$.*

The theory of images of polynomials on algebras has strong connections with the theory of polynomial identities (PI-theory). For instance, a polynomial identity for an algebra A is a polynomial whose image is $\{0\}$, and a central polynomial is a polynomial whose image lies in $Z(A)$, the center of the algebra A. The set of all polynomial identities of an algebra A is an ideal of $K\langle x_1, x_2, \ldots\rangle$ which is invariant under endomorphisms of $K\langle x_1, x_2, \ldots\rangle$. It is denoted by $T(A)$.

Some techniques from PI-theory are useful in studying images of polynomials on algebras. For instance, when $A = M_n(K)$, the m-generated algebra of generic matrices is known to be isomorphic to the quotient algebra

$$
F_m(M_n(K)) = \frac{K\langle x_1, \ldots, x_m\rangle}{T(M_n(K)) \cap K\langle x_1, \ldots, x_m\rangle},
$$

and working module $T(M_n(K))$ is equivalent to working in the algebra of generic matrices, see for instance [6, Chapter 7].

Let us denote by St_k the standard polynomial of degre k:

$$
St_k(x_1, \ldots, x_k) = \sum_{\sigma \in S_k} (-1)^\sigma x_{\sigma(1)} \cdots x_{\sigma(k)}.
$$

The following is a well-known fact about identities in matrices.

Theorem 2.3 (Amitsur–Levitzki) *The algebra $M_n(K)$ satisfies the polynomial identity St_{2n}. Moreover, $M_n(K)$ does not satisfy any polynomial identity of degree less than $2n$ and any polynomial identity of degree $2n$ of $M_n(K)$ is a scalar multiple of St_{2n}.*

Corollary 2.4 *The algebra $M_n(K)$ does not have central polynomials of degree less than $2n$.*

Proof Assume $f(x_1, \ldots, x_m)$ is a multilinear central polynomial for $M_n(K)$. Then the commutator $g = [f(x_1, \ldots, x_m), x_{m+1}]$ is a polynomial identity for $M_n(K)$. As a consequence $m + 1 \geq 2n$, which means $m \geq 2n - 1$. If $m = 2n - 1$ then g is a polynomial identity of degree $2n$, and it must be a scalar multiple of St_{2n}, but writing g as a sum of nonzero monomials gives us at most $2(2n-1)!$ summands, while in St_{2n} we have $(2n)!$ summands. A contradiction. Hence $m \geq 2n$. □

3. A new bound for the Mesyan conjecture

The main goal of this section is to present evidence which will allow us to state the Mesyan conjecture in a stronger form.

Let us assume for a moment that the Lvov–Kaplansky conjecture is true. If f is a polynomial of degree $m \leq 2n-1$ then by Corollary 2.4, f is not a central polynomial nor an identity for $M_n(K)$. By our assumption $f(M_n(K))$ is $sl_n(K)$ or $M_n(K)$ which, in both cases, contains $sl_n(K)$.

The above fact suggests that the Mesyan conjecture could be stated in a stronger form, namely for $n \geq \frac{m+1}{2}$. Also, this bound cannot be improved since St_{2n} is a polynomial identity of degree $2n$ for $M_n(K)$.

We now present more evidence that the conjecture should be stated in this form. We will prove a more general version of Proposition 1.2:

Theorem 3.1 *Let K be a field, $n \geq 2$ and $m \geq 2$ positive integers and let $f(x_1, \cdots, x_m)$ be a nonzero multilinear polynomial in $K\langle x_1, x_2, \ldots, x_m \rangle$. If $n \geq \frac{m+1}{2}$, then the K-subspace $span(f(M_n(K))$ contains $sl_n(K)$.*

Recall that for polynomials of degree $m = 2$, the Lvov-Kaplanksy conjecture is a consequence of Proposition 2.2 (4) and results of Shoda [19] and Albert and Muckenhoupt [1]. Also, for $m \geq 3$ we have $\frac{m+1}{2} \leq m - 1$, which shows that the above is a generalization of Proposition 1.2.

The following is a particular case of [10, Theorem 5]:

Theorem 3.2 *If U is a Lie ideal of $M_n(K)$, then either U is contained in the center of $M_n(K)$, or else U contains $sl_n(K)$, except in the case that $n = 2$ and K has characteristic 2.*

We are now ready to present a proof of Theorem 3.1.

Proof (of Theorem 3.1) Let

$$f(x_1, \cdots, x_m) = \sum_{\sigma \in S_m} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)}.$$

We first recall that if $m = 2$ and f is a nonzero polynomial, the image of f is $sl_n(K)$ or $M_n(K)$. Therefore, we can assume $m \geq 3$.

Without loss of generality we may assume that $\alpha_{(1)} \neq 0$. Let $i, j \in \{1, \ldots, n\}$ such that $i \neq j$. We will consider two cases: when m is even and when m is odd.

Case 1: m is even.

Let $m = 2k$, with k integer and $k \geq 2$, then $n \geq k + \frac{1}{2}$. Since k and n are integers, then $n \geq k + 1$, therefore $n - 2 \geq k - 1$. So let $l_1, \ldots, l_{k-1}$ be $k - 1$ distinct elements in $\{1, \ldots, n\} - \{i, j\}$. Then we have

$$f(e_{ii}, e_{ij}, e_{jl_1}, e_{l_1 l_1}, e_{l_1 l_2}, e_{l_2 l_2}, \ldots, e_{l_{k-2} l_{k-1}}, e_{l_{k-1} j}) = \alpha_{(1)} e_{ij}.$$

Since $\alpha_{(1)} \neq 0$ and e_{ij} with i, j distinct is a matrix whose diagonal contains only zeros, we have $span(f(M_n(K))$ contains all matrices with zeros on the main diagonal.

Case 2: m is odd.

Let $m = 2k - 1$, where k is an integer, such that $k \geq 2$. Since $n \geq \frac{m+1}{2}$ then $n - 2 \geq k - 2$. Therefore, we can find $k - 2$ distinct elements in $\{1, \ldots, n\} - \{i, j\}$, which we will denote by $l_1, \ldots, l_{k-2}$. Hence,

$$f(e_{ii}, e_{ij}, e_{jl_1}, e_{l_1 l_1}, e_{l_1 l_2}, e_{l_2 l_2}, \ldots, e_{l_{k-3} l_{k-2}}, e_{l_{k-2} l_{k-2}}, e_{l_{k-2} j}) = \alpha_{(1)} e_{ij}.$$

Therefore, we have proved that $span(f(M_n(K)))$ contains all matrices with zero diagonals. Assuming that $n \neq 2$ and $char(K) \neq 2$, and by Proposition 2.2 (3), we can apply Theorem 3.2, which implies that $span(f(M_n(K))) \supset sl_n(K)$.

Now we consider the case $M_2(K)$, where K is a field of characteristic 2. By hypothesis, we must have $m = 3$. Write

$$f(x_1, x_2, x_3) = \alpha_1 x_1 x_2 x_3 + \alpha_2 x_1 x_3 x_2 + \alpha_3 x_2 x_1 x_3 + \alpha_4 x_2 x_3 x_1 + \alpha_5 x_3 x_1 x_2 + \alpha_6 x_3 x_2 x_1$$

for some $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6 \in K$, with $\alpha_1 \neq 0$. Then,

$$f(e_{12}, e_{22}, e_{21}) = \alpha_1 e_{11} + \alpha_4 e_{22} + \alpha_5 e_{22}$$

and

$$f(e_{21}, e_{11}, e_{12}) = \alpha_1 e_{22} + \alpha_4 e_{11} + \alpha_5 e_{11}.$$

Subtracting the second equation from the first, gives $(\alpha_1 - \alpha_4 - \alpha_5)(e_{11} - e_{22}) \in span(f(M_2(K)))$. If $\alpha_1 - \alpha_4 - \alpha_5 \neq 0$, then this shows that $e_{11} - e_{22} \in span(f(M_2(K)))$. While if $\alpha_1 - \alpha_4 - \alpha_5 = 0$, then

$$f(e_{12}, e_{22}, e_{21}) = \alpha_1 e_{11} + \alpha_4 e_{22} + \alpha_5 e_{22} = \alpha_1(e_{11} + e_{22}) = \alpha_1(e_{11} - e_{22}),$$

since $char(K) = 2$, again showing that $e_{11} - e_{22} \in span(f(M_2(K)))$.

Noting that

$$f(e_{11}, e_{12}, e_{22}) = \alpha_1 e_{12}$$

and

$$f(e_{22}, e_{21}, e_{11}) = \alpha_1 e_{21},$$

from before, we conclude that $span(f(M_2(K))) \supset sl_2(K)$. □

Now we restate Mesyan conjecture in light of Theorem 3.1 and of the discussion of the beginning of this section.

Conjecture 3.3 (Mesyan conjecture restated) *Let K be a field, $n \geq 2$ and $m \geq 1$ be integers, and let $f(x_1, \ldots, x_m)$ be a nonzero multilinear polynomial in $K\langle x_1, \ldots, x_m \rangle$. If $m \leq 2n - 1$, then $sl_n(K) \subseteq f(M_n(K))$.*

4. The Mesyan conjecture for polynomials of degree 4

In this last section our main goal is to discuss the following result given in [5] and also to give a correction of a particular lemma used in its proof.

Theorem 4.1 *Let $n \geq 3$ and let K be an algebraically closed field of characteristic zero. Then the image of a nonzero multilinear polynomial $f(x_1, x_2, x_3, x_4)$ on the matrix algebra $M_n(K)$ contains $sl_n(K)$.*

For the sake of completeness we will present the proof of Theorem 4.1 in next. However some preliminary lemmas will be required, and we will present them below without their proofs. The first one is given in [5, Lemma 6].

Lemma 4.2 *Let $n \geq 3$ be an integer, let K be a field of characteristic zero and let $a_{i,j} \in K$ such that $\sum_{i=1}^{n} a_{i,i} = 0$. Then there exist $A, B, C \in M_n(K)$ such that $[A, [[A, B], [A, C]]] = \sum_{i=1}^{n} a_{i,i}e_{i,i} + \sum_{i=1}^{n-1} a_{i,i+1}e_{i,i+1}$.*

For the next two lemmas, see [2, Lemma 1.2, Proposition 1.8].

Lemma 4.3 *Let K be any field and let $A \in M_n(K)$ be a diagonal matrix with pairwise different entries in the main diagonal. Then $[A, M_n(K)]$ is the set of matrices whose diagonals entries are all 0.*

Lemma 4.4 *Let D be a division ring, $n \geq 2$ an integer, and $A \in M_n(D)$ noncentral matrix. Then, A is similar to a matrix in $M_n(D)$ with at most one nonzero entry on the main diagonal. In particular, if A has trace zero, then it is similar to a matrix in $M_n(D)$ with only zeros on the main diagonal.*

Proof [Proof of Theorem 4.1] The proof is the same as the one given in [5, Theorem 1]. We start the proof by reducing the polynomial f to a proper one. This can be done by considering the degree three multilinear polynomials obtained from f through the evaluation of some variable $x_i, i = 1, \ldots, 4$, by 1. In case one of these four polynomials is nonzero then we are able to use Mesyan's result (see [18, Theorem 13]) to obtain the desired conclusion. Otherwise, since $char(K) = 0$ then we conclude that f is as a proper polynomial (see for instance [6, Exercise 4.3.6]).

Hence we may write f as

$$f(x_1, x_2, x_3, x_4) = L(x_1, x_2, x_3, x_4) + \alpha_1[x_1, x_2][x_3, x_4] + \alpha_2[x_1, x_3][x_2, x_4]$$
$$+\alpha_3[x_1, x_4][x_2, x_3] + \alpha_4[x_2, x_3][x_1, x_4] + \alpha_5[x_2, x_4][x_1, x_3] + \alpha_6[x_3, x_4][x_1, x_2]$$

where $\alpha_1, \ldots, \alpha_6 \in K$. Using a Hall basis for the multilinear Lie polynomials of degree four (see [4, subsection 2.3]), we may write the Lie polynomial L as

$$L(x_1, x_2, x_3, x_4) = \beta_1[[[x_2, x_1], x_3], x_4] + \beta_2[[[x_3, x_1], x_2], x_4] + \beta_3[[[x_4, x_1], x_2], x_3]$$
$$+\beta_4[[x_4, x_1], [x_3, x_2]] + \beta_5[[x_4, x_2], [x_3, x_1]] + \beta_6[[x_4, x_3], [x_2, x_1]]$$

for some scalars $\beta_1, \ldots, \beta_6 \in K$. Since the three last terms of the Lie polynomial L can be written as a linear

combination of products of two commutators we may write f as

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4) =& \beta_1[[[x_2, x_1], x_3], x_4] + \beta_2[[[x_3, x_1], x_2], x_4] + \beta_3[[[x_4, x_1], x_2], x_3] \\
&+ \alpha_1[x_1, x_2][x_3, x_4] + \alpha_2[x_1, x_3][x_2, x_4] + \alpha_3[x_1, x_4][x_2, x_3] \\
&+ \alpha_4[x_2, x_3][x_1, x_4] + \alpha_5[x_2, x_4][x_1, x_3] + \alpha_6[x_3, x_4][x_1, x_2].
\end{aligned}
$$

If some among the scalars $\beta_1, \beta_2, \beta_3$ is nonzero, we claim that $sl_n(K) \subset f(M_n(K))$. Indeed, say $\beta_1 \neq 0$ without loss of generality. Take $x_1 = x_3 = x_4 = D$ as a diagonal matrix with pairwise distinct diagonal entries. Since any matrix in $M_n(K)$ is the sum of a diagonal matrix and a matrix with zeros in the main diagonal, Lemma 4.3 implies that $f(D, x_2, D, D)$ consists of all matrices with zeros in the main diagonal. On the other hand, Lemma 4.4 states that traceless matrices are conjugate to those matrices with zero diagonal. Then by Proposition 2.2 (1) the claim is proved. The cases where $\beta_2 \neq 0$ and $\beta_3 \neq 0$ can be handled similarly.

From now on we may assume $\beta_1 = \beta_2 = \beta_3 = 0$ and then we consider the two following cases.

Case 1: $\alpha_1 = \alpha_4 = \alpha_6 = \alpha_3 = -\alpha_2 = -\alpha_5$.

The above assumptions on the coefficients of f lead us to $f = \lambda St_4$ where $\lambda \in K \setminus \{0\}$. Using the identity $[uv, w] = [u, w]v + u[v, w]$ we have

$$
\begin{aligned}
St_4(A, A^2, B, C) &= [A, A^2][B, C] + [B, C][A, A^2] + [A^2, B][A, C] + [A, C][A^2, B] - [A, B][A^2, C] - [A^2, C][A, B] \\
&= [A^2, B][A, C] + [A, C][A^2, B] - [A, B][A^2, C] - [A^2, C][A, B]. \\
&= [A, [[A, B], [A, C]]]
\end{aligned}
$$

Now it is enough to apply Lemma 4.2 for the Jordan normal form of a traceless matrix.

Case 2: *At least one among the equalities $\alpha_1 = \alpha_4 = \alpha_6 = \alpha_3 = -\alpha_2 = -\alpha_5$ does not hold.*

In this case one may check that there exist matrices $A, B, C \in M_n(K)$ such that at least one of the following is a nonzero matrix

$$
\begin{aligned}
f(A, A, B, C) &= (\alpha_2 + \alpha_4)[A, B][A, C] + (\alpha_3 + \alpha_5)[A, C][A, B] \\
f(A, B, A, C) &= (\alpha_1 - \alpha_4)[A, B][A, C] + (\alpha_6 - \alpha_3)[A, C][A, B] \\
f(A, B, C, A) &= (-\alpha_1 - \alpha_5)[A, B][A, C] + (-\alpha_2 - \alpha_6)[A, C][A, B] \\
f(B, A, A, C) &= (-\alpha_1 - \alpha_2)[A, B][A, C] + (-\alpha_5 - \alpha_6)[A, C][A, B] \\
f(B, A, C, A) &= (-\alpha_3 + \alpha_1)[A, B][A, C] + (\alpha_6 - \alpha_4)[A, C][A, B] \\
f(B, C, A, A) &= (\alpha_2 + \alpha_3)[A, B][A, C] + (\alpha_4 + \alpha_5)[A, C][A, B].
\end{aligned}
$$

Hence it is enough to study the image of the polynomial

$$
f = [x_1, x_2][x_1, x_3] + \lambda[x_1, x_3][x_1, x_2] \tag{4.1}
$$

on $M_n(K)$, where $\lambda \in K$. This will be done in the next two lemmas. $\qquad\square$

The next lemma is [5, Lemma 5].

Lemma 4.5 *Let K be an algebraically closed field of characteristic zero and let $n \geq 3$. Then each $D \in sl_n(K)$ can be written as $D = [[A, B], [A, C]]$ for a suitable choice of matrices $A, B, C \in M_n(K)$.*

We note that Lemma 4.5 completely solves the case where $\lambda = -1$ in (4.1). The next lemma is [5, Lemma 4] and it deals with the others values for λ.

Lemma 4.6 *Let K be an algebraically closed field of characteristic zero, let $n \geq 3$ and let $\lambda \in K \setminus \{-1\}$. Then each $D \in sl_n(K)$ can be written as $D = [A, B][A, C] + \lambda[A, C][A, B]$ for a suitable choice of matrices $A, B, C \in M_n(K)$.*

We first note that the proof of Lemma 4.6 presented in [5] is not correct, since it was claimed that given scalars $d_1, \ldots, d_n \in K$ satisfying $\sum_{i=1}^{n} d_i = 0$ the following system of equations

$$
\begin{cases}
(\lambda + 1)y_1 & = & d_1 \\
(\lambda + 1)y_2 & = & d_2 \\
& \vdots & \\
(\lambda + 1)y_{n-1} & = & d_{n-1} \\
-(n-1)(\lambda + 1)y_n & = & d_n
\end{cases}
$$

has a solution $y_i = b_i, i = 1, \ldots, n$ satisfying $\sum_{i=1}^{n} b_i = 0$. However the existence of such solution would give us $-n(\lambda + 1)b_n = 0$ by summing all equations above. Hence $b_n = 0$, which implies $d_n = 0$, contradicting the generality of the chosen d_n.

Next we will present a correction of the proof of Lemma 4.6. We recall the following lemma from [5, Lemma 2] which will be used in our proof.

Lemma 4.7 *Let K be a field, let $a_{i,j} \in K$ such that $\sum_{i=1}^{n} a_{i,i} = 0$ and let $A = \sum_{i=1}^{n-1} e_{i,i+1} \in M_n(K)$. Then there exists $B \in M_n(K)$ such that*

$$
[A, B] = \sum_{i=1}^{n} a_{i,i}e_{i,i} + \sum_{i=1}^{n-1} a_{i,i+1}e_{i,i+1}.
$$

Proof [Proof of Lemma 4.6] We start noting that we may assume D is in its Jordan normal form since K is an algebraically closed field, and the image of the polynomial

$$
f(x_1, x_2, x_3) = [x_1, x_2][x_1, x_3] + \lambda[x_1, x_3][x_1, x_2]
$$

is closed under conjugation by invertible elements of $M_n(K)$.

So we write D as

$$
D = \sum_{i=1}^{n} d_{ii}e_{ii} + \sum_{i=1}^{n-1} d_{i,i+1}e_{i,i+1}, \tag{4.2}
$$

where $d_{ii}, d_{i,i+1} \in K$. Take $A = \sum_{i=1}^{n-1} e_{i,i+1}$, and consider any $a_{ii}, b_{ii} \in K$, $i = 1, \ldots, n$ and $b_{i,i+1} \in K$, $i = 1, \ldots, n-1$, such that

$$
\sum_{i=1}^{n} a_{ii} = 0 = \sum_{i=1}^{n} b_{ii}.
$$

Lemma 4.7 gives us the existence of matrices $B, C \in M_n(K)$ where

$$[A, B] = \sum_{i=1}^{n} a_{ii} e_{ii} \text{ and } [A, C] = \sum_{i=1}^{n} b_{ii} e_{ii} + \sum_{i=1}^{n-1} b_{i,i+1} e_{i,i+1}.$$

Therefore,

$$[A, B][A, C] + \lambda[A, C][A, B] = (1 + \lambda) \sum_{i=1}^{n} a_{ii} b_{ii} e_{ii} + \sum_{i=1}^{n-1} (a_{ii} + \lambda a_{i+1,i+1}) b_{i,i+1} e_{i,i+1} \tag{4.3}$$

and we are looking for a simultaneous solution of the two systems below given by comparing Equations (4.2) and (4.3)

$$\begin{cases} (1 + \lambda)a_{11}b_{11} = d_{11} \\ \qquad \vdots \\ (1 + \lambda)a_{nn}b_{nn} = d_{nn} \end{cases} \tag{4.4}$$

and

$$\begin{cases} (a_{11} + \lambda a_{22})b_{12} = d_{12} \\ \qquad \vdots \\ (a_{n-1,n-1} + \lambda a_{nn})b_{n-1,n} = d_{n-1,n} \end{cases} \tag{4.5}$$

jointly with the conditions $\sum_{i=1}^{n} a_{ii} = 0 = \sum_{i=1}^{n} b_{ii}$.

Now we will divide our proof into the following three cases depending on the number of Jordan blocks in the Jordan normal form of D.

Case 1: *D has exactly one Jordan block.*

In this case we must have $d_{ii} = 0$ for all i, since D is a traceless matrix and $char(K) = 0$.

If $\lambda \neq \dfrac{1}{n-1}$, one can check that is enough to choose $b_{11} = \cdots = b_{nn} = 0, a_{11} = \cdots = a_{n-1,n-1} = 1, a_{nn} = -(n-1), b_{12} = \dfrac{d_{12}}{1+\lambda}, b_{23} = \dfrac{d_{23}}{1+\lambda}, \cdots, b_{n-2,n-1} = \dfrac{d_{n-2,n-1}}{1+\lambda}$ and $b_{n-1,n} = \dfrac{d_{n-1,n}}{1+\lambda(1-n)}$.

If $\lambda = \dfrac{1}{n-1} \neq 0$, then we choose $b_{11} = \cdots = b_{nn} = 0, a_{11} = \cdots = a_{n-2,n-2} = 1, a_{n-1,n-1} = 0, a_{nn} = -(n-2), b_{12} = \dfrac{d_{12}}{1+\lambda}, \cdots, b_{n-3,n-2} = \dfrac{d_{n-3,n-2}}{1+\lambda}, b_{n-2,n-1} = d_{n-2,n-1}$ and $b_{n-1,n} = \dfrac{d_{n-1,n}}{\lambda(2-n)}$, and we are done with the first case.

Assuming $a_{ii} \neq 0$, $i = 1, \ldots, n$, by the equations in (4.4) we have $(1 + \lambda)a_{ii}b_{ii} = d_{ii}$ for $i = 1, \ldots, n$ and then $b_{ii} = (1 + \lambda)^{-1} a_{ii}^{-1} d_{ii}$. Summing these equations for all i we get that $(1 + \lambda)^{-1} (\sum_{i=1}^{n} a_{ii}^{-1} d_{ii}) = 0$, that is,

$$\frac{d_{11}}{a_{11}} + \cdots + \frac{d_{nn}}{a_{nn}} = 0.$$

Case 2: *D has exactly two Jordan blocks.*

Suppose D has one block of size m_1 and eigenvalue d_1, and another block of size m_2 and eigenvalue d_2. One of these two blocks must be of size at least two, since $n \geq 3$, and therefore we can take $m_1 \geq 2$. Since D has trace zero we have $d_1 = 0$ if and only if $d_2 = 0$, and then the previous case allow us to assume that $d_1 \neq 0$. In this case we are looking for nonzero values for all a_{ii} such that $a_{ii} + \lambda a_{i+1,i+1}$ is nonzero for all i. Note that this last condition can be used to compute the values of $b_{i,i+1}$ in (4.5) easily.

Taking $a_{33} = \cdots = a_{nn} = 1$, we obtain

$$0 = \frac{d_1}{a_{11}} + \cdots + \frac{d_1}{a_{m_1,m_1}} + \frac{d_2}{a_{m_1+1,m_1+1}} + \cdots + \frac{d_2}{a_{nn}} = \frac{d_1}{a_{11}} + \frac{d_1}{-a_{11} - (n-2)} + (m_1 - 2)d_1 - m_1 d_1.$$

Therefore

$$\frac{1}{a_{11}} - \frac{1}{a_{11} + (n-2)} - 2 = 0, \text{ i.e. } a_{11} + (n-2) - a_{11} - 2a_{11}^2 - 2(n-2)a_{11} = 0,$$

which leads us to the following equation

$$2a_{11}^2 + 2(n-2)a_{11} - (n-2) = 0. \tag{4.6}$$

Since $n - 2 \neq 0$, then $a_{11} \neq 0$. We have also $a_{11} \neq -(n-2)$, since otherwise

$$2(n-2)^2 + 2(n-2)(-(n-2)) - (n-2) = 0, \text{ that is, } n = 2.$$

We note that the roots of Equation (4.6) in a_{11} are

$$a_{11} = \frac{-(n-2) \pm \sqrt{n(n-2)}}{2}$$

and so

$$a_{22} = \frac{-(n-2) \mp \sqrt{n(n-2)}}{2},$$

provided that $a_{22} = -a_{11} - (n-2)$.

In computing $b_{i,i+1}$, the variables b_{12} and b_{23} depend on a_{11} and a_{22} via the equations

$$(a_{11} + \lambda a_{22})b_{12} = d_{12} \text{ and } (a_{22} + \lambda)b_{23} = d_{23}.$$

Denote

$$\bar{a}_1 = \frac{-(n-2) + \sqrt{n(n-2)}}{2} \text{ and } \bar{a}_2 = \frac{-(n-2) - \sqrt{n(n-2)}}{2}.$$

For $\lambda \neq -\bar{a}_1/\bar{a}_2$ and $\lambda \neq -\bar{a}_2$, we take $a_{11} = \bar{a}_1$ and $a_{22} = \bar{a}_2$. For $\lambda \neq -\bar{a}_1/\bar{a}_2$ and $\lambda = -\bar{a}_2$, we take $a_{11} = \bar{a}_2$ and $a_{22} = \bar{a}_1$. So in the first equation we will have $\bar{a}_2 - \bar{a}_2\bar{a}_1$, which is nonzero since $\bar{a}_1 \neq 1$, and in the second equation we will have $\bar{a}_1 - \bar{a}_2$ that is also nonzero. The last case is $\lambda = -\bar{a}_1/\bar{a}_2$. Again we take

$a_{11} = \bar{a}_2$ and $a_{22} = \bar{a}_1$. Hence $\bar{a}_2 + \lambda \bar{a}_1 \neq 0$, since otherwise $\bar{a}_1/\bar{a}_2 = \bar{a}_2/\bar{a}_1$, i.e. $\bar{a}_1^2 = \bar{a}_2^2$ which is an absurd. We also have $\bar{a}_1 - \bar{a}_1/\bar{a}_2 \neq 0$, since otherwise $\bar{a}_2 = 1$, which is likewise absurd.

Since all $\lambda \in K \setminus \{-1\}$ was considered above we finished the proof of the second case.

Case 3: D has $k \geq 3$ Jordan blocks.

Now suppose that D is in the Jordan normal form with $k \geq 3$. We denote by m_l the size of the l-th Jordan block of D. For the matrix $[A, B] = \sum_{i=1}^{n} a_{ii} e_{ii}$, we will consider the same block division that occurs in D, and in a same block we take all a_{ii} equal to each other. For every $j \in \{1, \ldots, k\}$, we will denote the element on the main diagonal of the j-th block of $[A, B]$ by a_j.

We assume that $a_j \neq 0$ for all $j \in \{1, \ldots, k\}$. Since $(1 + \lambda) a_{ii} b_{ii} = d_{ii}$, then $b_{ii} = (1 + \lambda)^{-1} a_{ii}^{-1} d_{ii}$ and so $\sum_{i=1}^{n} a_{ii}^{-1} d_{ii} = 0$. Therefore,

$$0 = \frac{m_1 d_1}{a_1} + \cdots + \frac{m_{k-1} d_{k-1}}{a_{k-1}} + \frac{m_k d_k}{a_k}$$

$$= \frac{m_1 d_1}{a_1} + \cdots + \frac{m_{k-1} d_{k-1}}{a_{k-1}} + \frac{-\sum_{j=1}^{k-1} m_j d_j}{-\sum_{j=1}^{k-1} \frac{m_j}{m_k} a_j}$$

$$= \frac{m_1 d_1}{a_1} + \cdots + \frac{m_{k-1} d_{k-1}}{a_{k-1}} + \frac{m_k \sum_{j=1}^{k-1} m_j d_j}{\sum_{j=1}^{k-1} m_j a_j}.$$

Taking $a_1 = \cdots = a_{k-2} = 1$, we obtain

$$0 = m_1 d_1 + \cdots + m_{k-2} d_{k-2} + \frac{m_{k-1} d_{k-1}}{a_{k-1}} + \frac{m_k \sum_{j=1}^{k-1} m_j d_j}{\sum_{j=1}^{k-2} m_j + m_{k-1} a_{k-1}},$$

and hence

$$0 = a_{k-1} \left(\sum_{j=1}^{k-2} m_j + m_{k-1} a_{k-1} \right) \left(\sum_{j=1}^{k-2} m_j d_j \right) + m_{k-1} d_{k-1} \left(\sum_{j=1}^{k-2} m_j + m_{k-1} a_{k-1} \right) + a_{k-1} \left(m_k \sum_{j=1}^{k-1} m_j d_j \right)$$

$$= m_{k-1} \left(\sum_{j=1}^{k-2} m_j d_j \right) a_{k-1}^2 + \left(\left(\sum_{j=1}^{k-2} m_j \right) \left(\sum_{j=1}^{k-2} m_j d_j \right) + m_{k-1}^2 d_{k-1} + m_k \sum_{j=1}^{k-1} m_j d_j \right) a_{k-1} + m_{k-1} d_{k-1} \sum_{j=1}^{k-2} m_j.$$

Setting $d = \sum_{j=1}^{k-2} m_j d_j$ we obtain the follow quadratic equation in a_{k-1}:

$$m_{k-1} d a_{k-1}^2 + \left(d \sum_{j=1}^{k-2} m_j + m_k d + m_{k-1} d_{k-1}(m_{k-1} + m_k) \right) a_{k-1} + m_{k-1} d_{k-1} \sum_{j=1}^{k-2} m_j = 0. \qquad (4.7)$$

We want that $a_{k-1} \neq 0$ and $a_k \neq 0$. Since $a_k = \dfrac{-\sum_{j=1}^{k-2} m_j - m_{k-1} a_{k-1}}{m_k}$, we are looking for some nonzero solution of (4.7) different from $-\dfrac{\sum_{j=1}^{k-2} m_j}{m_{k-1}}$. We divide this task in the three following subcases.

Subcase 1: $m_{k-1} d = 0$.

In this last case we have $d = 0$ and so Equation (4.7) turns into

$$d_{k-1} \left((m_{k-1} + m_k) a_{k-1} + \sum_{j=1}^{k-2} m_j \right) = 0.$$

If $d_{k-1} = 0$, then any element of K is solution and therefore we choose an appropriate one.

If $d_{k-1} \neq 0$, then $a_{k-1} = -\dfrac{\sum_{j=1}^{k-2} m_j}{m_{k-1} + m_k}$. Given that $k \geq 3$ we have $a_{k-1} \neq 0$, and since $m_k \neq 0$ we obtain $a_{k-1} \neq -\sum_{j=1}^{k-2} \dfrac{m_j}{m_{k-1}}$.

Subcase 2: $m_{k-1} d \neq 0$ and $m_{k-1} d_{k-1} \sum_{j=1}^{k-2} m_j = 0$.

Since m_{k-1} and $\sum_{j=1}^{k-2} m_j$ are nonzero, then $d_{k-1} = 0$. So Equation (4.7) can be rewritten as

$$m_{k-1} d a_{k-1}^2 + d \left(\sum_{j=1}^{k-2} m_j + m_k \right) a_{k-1} = 0.$$

A solution of the equation above is $a_{k-1} = -\dfrac{\sum_{j=1}^{k-2} m_j + m_k}{m_{k-1}} \neq 0$ which is also different from $-\sum_{j=1}^{k-2} \dfrac{m_j}{m_{k-1}}$, since otherwise we would have $m_k = 0$, a contradiction.

Subcase 3: $m_{k-1} d \neq 0$ and $m_{k-1} d_{k-1} \sum_{j=1}^{k-2} m_j \neq 0$.

We have necessarily two nonzero solutions in this case. Now we prove that at least one of them is different from $-\dfrac{\sum_{j=1}^{k-2} m_j}{m_{k-1}}$. Suppose, seeking a contradiction, that Equation (4.7) has two repeated roots equal

to $-\sum_{j=1}^{k-2} \dfrac{m_j}{m_{k-1}}$, i.e.

$$\left(a_{k-1} + \sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}} \right)^2 = 0. \tag{4.8}$$

Hence $a_{k-1} = -\sum_{j=1}^{k-2} \dfrac{m_j}{m_{k-1}}$ and, on other hand, using the well known formula for the sum of the roots of a quadratic equation and recalling that we have two repeated roots, we also have

$$a_{k-1} = -\frac{d\sum_{j=1}^{k-2} m_j + m_k d + m_{k-1} d_{k-1}(m_{k-1}+m_k)}{2m_{k-1}d} = -\sum_{j=1}^{k-2} \frac{m_j}{2m_{k-1}} - \frac{m_k}{2m_{k-1}} - \frac{d_{k-1}}{2d}(m_{k-1}+m_k).$$

Therefore,

$$\frac{d_{k-1}}{2d}(m_{k-1}+m_k) = -\frac{m_k}{2m_{k-1}} + \frac{1}{2}\sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}}, \text{i.e.}$$

$$\frac{d_{k-1}}{d} = \frac{\sum_{j=1}^{k-2} m_j - m_k}{m_{k-1}(m_{k-1}+m_k)}. \tag{4.9}$$

Dividing Equation (4.7) by $m_{k-1}d$ and comparing with Equation (4.8), we obtain

$$\begin{cases} \displaystyle\sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}} + \frac{m_k}{m_{k-1}} + \frac{d_{k-1}}{d}(m_{k-1}+m_k) = 2\sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}} \\[2ex] \displaystyle\frac{d_{k-1}}{d}\sum_{j=1}^{k-2} m_j = \frac{1}{m_{k-1}^2}\left(\sum_{j=1}^{k-2} m_j\right)^2. \end{cases}$$

Since $\sum_{j=1}^{k-2} m_j \neq 0$, then $\dfrac{d_{k-1}}{d} = \dfrac{1}{m_{k-1}^2}\left(\sum_{j=1}^{k-2} m_j\right)$ and so we get

$$2m_{k-1}\frac{d_{k-1}}{d} = \sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}} + \frac{m_k}{m_{k-1}} + \frac{d_{k-1}}{d}(m_{k-1}+m_k)$$

$$= \frac{n-m_{k-1}}{m_{k-1}} + \frac{d_{k-1}}{d}(m_{k-1}+m_k),$$

which implies that

$$\frac{d_{k-1}}{d}(m_{k-1}-m_k) = \frac{n-m_{k-1}}{m_{k-1}}.$$

By Equation (4.9), we have

$$\frac{n - m_{k-1}}{m_{k-1}} = \frac{\sum_{j=1}^{k-2} m_j - m_k}{m_{k-1}(m_{k-1} + m_k)}(m_{k-1} - m_k), \text{ that is,}$$

$$n - m_{k-1} = \frac{n - m_{k-1} - 2m_k}{m_{k-1} + m_k}(m_{k-1} - m_k).$$

Therefore we have

$$(n - m_{k-1})(m_{k-1} + m_k) = (n - m_{k-1} - 2m_k)(m_{k-1} - m_k),$$

and expanding the brackets we obtain

$$nm_{k-1} + nm_k - m_{k-1}^2 - m_{k-1}m_k = nm_{k-1} - nm_k - m_{k-1}^2 + m_{k-1}m_k - 2m_{k-1}m_k + 2m_k^2$$

which implies that

$$m_k(n - m_k) = 0, \text{ which means } m_k = 0 \text{ ou } n = m_k,$$

and in the both cases we get a contradiction.

We conclude that there exists a nonzero root of (4.7) different from $-\sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}}$.

Now it is enough to determine the values for $b_{i,i+1}$ in the system (4.5).

In the matrix D, from the end of the block that contains the element d_{ii} to the beginning of the one containing $d_{i+1,i+1}$ we have $d_{i,i+1} = 0$, and then we can take $b_{i,i+1} = 0$. For the other elements above the main diagonal we have $(a_{ll} + \lambda a_{l+1,l+1})b_{l,l+1} = d_{l,l+1}$ and $a_{ll} = a_{l+1,l+1}$. So we can take $b_{l,l+1} = a_{ll}^{-1}(1 + \lambda)^{-1}d_{l,l+1}$.

This finishes the proof of the third and last case, and therefore we conclude the proof of the lemma. □

References

[1] Albert AA, Muckenhoupt B. On matrices of trace zeros. Michigan Mathematical Journal 1957; 4 (1): 1-3. doi: 10.1307/mmj/1028990168

[2] Amitsur S, Rowen L. Elements of reduce trace 0. Israel Journal of Mathematics 1994; 87: 161-179. doi: 10.1007/BF02772992

[3] Anzis BE, Emrich ZM, Valiveti KG. On the images of Lie polynomials evaluated on Lie algebras. Linear Algebra and its Applications 2015; 469:51-75. doi: 10.1016/j.laa.2014.11.015

[4] Bahturin YA. Identical relations in Lie algebras. VNU Science Press, b.v., Utrecht, 1987. Translated from the Russian by Bahturin.

[5] Buzinski D, Winstanley R. On multilinear polynomials in four variables evaluated on matrices. Linear Algebra and its Applications 2013; 439 (9): 2712-2719. doi: 10.1016/j.laa.2013.08.005

[6] Drensky V. Free algebras and PI-algebras. Springer-Verlag Singapore, Singapore, 2000. Graduate course in algebra.

[7] Fagundes PS. The images of multilinear polynomials on strictly upper triangular matrices. Linear Algebra and its Applications 2019; 563: 287-301. doi: 10.1016/j.laa.2018.11.014

[8] Fagundes PS, de Mello, TC. Images of multilinear polynomials of degree up to four on upper triangular matrices. Operators and Matrices 2019; 13 (1): 283-292. doi: 10.7153/oam-2019-13-18

[9] Gargate IG, de Mello TC. Images of multilinear polynomials on $n \times n$ upper triangular matrices over infinite fields. Israel Journal of Mathematics, to appear.

[10] Herstein, IN, On the Lie and Jordan Rings of a Simple Associative Ring. American Journal of Mathematics 1955; 77 (2): 279-285. doi: 10.2307/2372531

[11] Kanel-Belov A, Malev S, Rowen L. The images of non-commutative polynomials evaluated on 2×2 matrices. Proceedings of the American Mathematical Society 2012; 140 (2): 465-478. doi: 10.1090/S0002-9939-2011-10963-8

[12] Kanel-Belov A, Malev S, Rowen L. The images of multilinear polynomials evaluated on 3×3 matrices. Proceedings of the American Mathematical Society 2016; 144 (1): 7-19. doi: 10.1090/proc/12478

[13] Kanel-Belov A, Malev S, Rowen L. The images of Lie polynomials evaluated on matrices. Communications in Algebra 2017; 45 (11): 4801-4808. doi: 10.1080/00927872.2017.1282959

[14] Kanel-Belov A, Malev S, Rowen L, Yavich R. Evaluations of noncommutative polynomials on algebras: methods and problems, and the L'vov-Kaplansky conjecture. SIGMA Symmetry, Integrability and Geometry. Methods and Applications 2020; 16:Paper No. 071, 61. doi: 10.3842/SIGMA.2020.071

[15] Malev S. The images of non-commutative polynomials evaluated on 2×2 matrices over an arbitrary field. Journal of Algebra and its Applications 2014; 13 (6): 1450004, 12. doi: 10.1142/S0219498814500042

[16] Malev S. The images of noncommutative polynomials evaluated on the quaternion algebra. Journal of Algebra and its Applications 2021; 20 (5): Paper No. 2150074, 8. doi: 10.1142/S0219498821500742

[17] Malev S, Yavich R, Shayer R. Evaluations of multilinear polynomials on low rank Jordan algebras 2021; arXiv:2107.05266.

[18] Mesyan Z. Polynomials of small degree evaluated on matrices. Linear and Multilinear Algebra 2013; 61 (11): 1487-1495. doi: 10.1080/03081087.2012.758262

[19] Shoda K. Einige Sätze über Matrizen. Japanese Journal of Mathematics 1937; 13 (3): 361-365. doi: 10.4099/jjm1924.13.0_361

[20] Vitas D. Images of multilinear polynomials in the algebra of finitary matrices contain trace zero matrices. Linear Algebra and its Applications 2021; 626: 221-233. doi: 10.1016/j.laa.2021.05.018

[21] Vitas D. Multilinear polynomials are surjective on algebras with surjective inner derivations. Journal of Algebra 2021; 565: 255-281. doi: 10.1016/j.jalgebra.2020.09.004