

Evaluations of Noncommutative Polynomials on Algebras: Methods and Problems, and the L’vov–Kaplansky Conjecture)

Alexei KANEL-BELOV[†], Sergey MALEV[‡], Louis ROWEN[§] and Roman YAVICH[‡]

[†] Bar-Ilan University, MIPT, Israel
E-mail: kanelster@gmail.com

[‡] Department of Mathematics, Ariel University of Samaria, Ariel, Israel
E-mail: sergeyma@ariel.ac.il, romany@ariel.ac.il

[§] Department of Mathematics, Bar Ilan University, Ramat Gan, Israel
E-mail: rowen@math.biu.ac.il

Received September 18, 2019, in final form July 08, 2020; Published online July 27, 2020

<https://doi.org/10.3842/SIGMA.2020.071>

Abstract. Let p be a polynomial in several non-commuting variables with coefficients in a field K of arbitrary characteristic. It has been conjectured that for any n , for p multilinear, the image of p evaluated on the set $M_n(K)$ of n by n matrices is either zero, or the set of scalar matrices, or the set $\text{sl}_n(K)$ of matrices of trace 0, or all of $M_n(K)$. This expository paper describes research on this problem and related areas. We discuss the solution of this conjecture for $n = 2$ in Section 2, some decisive results for $n = 3$ in Section 3, and partial information for $n \geq 3$ in Section 4, also for non-multilinear polynomials. In addition we consider the case of K not algebraically closed, and polynomials evaluated on other finite dimensional simple algebras (in particular the algebra of the quaternions). This review recollects results and technical material of our previous papers, as well as new results of other researches, and applies them in a new context. This article also explains the role of the Deligne trick, which is related to some nonassociative cases in new situations, underlying our earlier, more straightforward approach. We pose some problems for future generalizations and point out possible generalizations in the present state of art, and in the other hand providing counterexamples showing the boundaries of generalizations.

Key words: L’vov–Kaplansky conjecture; noncommutative polynomials; multilinear polynomial evaluations; power central polynomials; the Deligne trick; PI algebras

2020 Mathematics Subject Classification: 16H05; 16H99; 16K20; 16R30; 16R40; 17B99

On occasion of the 80-th birthday of D.B. Fuchs

Contents

1	Introduction	2
1.1	Evaluations of words	5
1.2	Non-multilinear polynomials	6
1.3	The main theorems	7
1.3.1	The main theorems for $n = 2$	7
1.3.2	The main theorems for $n = 3$	8
1.3.3	Higher n	9
1.3.4	Polynomials over quaternion algebras	9
1.4	Some basic tools	10

1.4.1	Generic matrices	11
1.4.2	Generic elements	13
1.4.3	Basic facts about cones	14
1.5	Some general observations	14
2	The low rank case	15
2.1	2×2 matrices	16
2.1.1	The case of a quadratically closed field	16
2.1.2	Images of arbitrary non-homogeneous polynomials	20
2.1.3	The cases of an arbitrary real field	21
2.1.4	Images of multilinear polynomials evaluated on $M_2(K)$	21
2.1.5	Multilinear polynomials evaluated on \mathbb{H}	25
2.1.6	Semi-homogeneous polynomials evaluated on \mathbb{H}	27
2.2	The Deligne trick	28
2.2.1	Evaluations of word maps on matrix groups	28
2.2.2	Evaluations of semihomogeneous polynomials on matrix rings	30
2.2.3	Evaluations of multilinear polynomials on matrix rings	30
2.3	Some important questions and open problems regarding evaluations of polynomials on low rank algebras	32
3	The case of matrices of rank 3	34
3.1	3×3 matrices over a field with a primitive cube root of 1	34
3.2	Multilinear trace vanishing polynomials	34
3.2.1	Multilinear trace vanishing polynomials over an algebraically closed field	36
3.3	Deligne trick for algebras of rank 3	37
3.3.1	Evaluations of semihomogeneous polynomials on 3×3 matrices	37
3.3.2	Evaluations of multilinear trace zero polynomials on 3×3 matrices	38
3.4	The Euler graph approach	39
3.5	Open problems related to the rank 3 case	41
4	The case of high rank	42
4.1	The Euler graph approach for algebras of high rank	42
4.1.1	The idea of the method	42
4.1.2	Applications of the method	42
4.1.3	Questions related to the Euler graph approach	46
4.2	Non-trivial images: power-central polynomials on matrices	46
4.2.1	Considerations arising from division algebras	47
4.2.2	2-central polynomials linear in the first indeterminate	49
4.3	Open problems	50
4.3.1	Images of polynomial maps and matrix equations	50
5	Lie polynomials and non-associative generalizations	51
5.1	Homogeneous Lie polynomials on \mathfrak{gl}_n and \mathfrak{sl}_n	51
5.2	The case $n = 2$	54
5.3	The case $n = 3$	55
5.3.1	A group theoretical question and its relation to the Lie theoretical problem	55
5.4	Open problems for nonassociative algebras	56
	References	57

1 Introduction

In this review, we present a systematized exposition including results obtained in previous papers. In addition, we have systematized ideas and methods of proofs. Surprisingly, for quite elementary looking results, we have used the Deligne trick, and a technique for working with central simple algebras dating back to Amitsur. At the end of each section, we present open

problems and our ideas, not all of which we were able to complete. We would be happy if other scientists will succeed.

$K\langle x_1, \dots, x_m \rangle$ denotes the free K -algebra generated by noncommuting variables x_1, \dots, x_m ; we refer to the elements of $K\langle x_1, \dots, x_m \rangle$ as *polynomials*. Consider any algebra R over a field K . A polynomial $p \in K\langle x_1, \dots, x_m \rangle$ is called a *polynomial identity* (PI) of the algebra R if $p(a_1, \dots, a_m) = 0$ for all $a_1, \dots, a_m \in R$; $p \in K\langle x_1, \dots, x_m \rangle$ is a *central polynomial* of R , if for any $a_1, \dots, a_m \in R$ one has $p(a_1, \dots, a_m) \in \text{Cent}(R)$ but p is not a PI of R .

For any polynomial $p \in K\langle x_1, \dots, x_m \rangle$, the *image* of p (in R) is defined as

$$\text{Im } p := \{r \in R : \text{there exist } a_1, \dots, a_m \in R \text{ such that } p(a_1, \dots, a_m) = r\}.$$

Remark 1.1. $\text{Im } p$ is invariant under conjugation, since

$$sp(x_1, \dots, x_m)s^{-1} = p(sx_1s^{-1}, sx_2s^{-1}, \dots, sx_ms^{-1}) \in \text{Im } p,$$

for any invertible element s .

Images of polynomials evaluated on algebras play an important role in noncommutative algebra. In particular, various challenging problems related to the theory of polynomial identities have been settled after the construction of central polynomials by Formanek [41] and Razmyslov [89].

This survey consists of the main results by the authors describing the possible images of polynomials, especially in connection with the partial solution of a conjecture attributed to L'vov and Kaplansky concerning evaluations of polynomials on matrices that was formulated in [29]:

Conjecture 1.2 (L'vov–Kaplansky). *Let p be a multilinear polynomial. Then the set of values of p on the matrix algebra $M_n(K)$ over an infinite field K is a vector space.*

Remark 1.3. It is not difficult to ascertain the linear span of the values of any multilinear polynomial. Indeed, the linear span of its values comprises a Lie ideal since, as is well known,

$$[a, p(a_1, \dots, a_n)] = p([a, a_1], a_2, \dots, a_n) + p(a_1, [a, a_2], \dots, a_n) + \dots + p(a_1, \dots, [a, a_n]),$$

and Herstein [52] characterized Lie ideals of a simple ring R as either being contained in the center or containing the commutator Lie ideal $[R, R]$. Another proof is given in [18]; also see [58, Lemma 4]. It is considerably more difficult to determine the actual image set $\text{Im } p$, rather than its linear span.

Thus Conjecture 1.2 is equivalent to the following:

Conjecture 1.4. *If p is a multilinear polynomial evaluated on the matrix ring $M_n(K)$, then $\text{Im } p$ is either $\{0\}$, K , $\text{sl}_n(K)$, or $M_n(K)$. Here K indicates the set of scalar matrices and $\text{sl}_n(K)$ is the set of matrices of trace zero.*

Note that these options are mutually exclusive when $\text{char}(K)$ does not divide n . While lacking a verification of this conjecture, one is led to the following more general question:

Question 1.5. *Given a polynomial p (not necessarily multilinear), what is its possible image set? Which polynomial of minimal degree produces one of these image sets?*

Recall the **standard polynomial**

$$s_k := \sum_{\pi \in S_k} \text{sgn}(\pi) x_{\pi(1)} \cdots x_{\pi(k)}.$$

A polynomial p is *trace vanishing* if each of its evaluations has trace 0. For example, it is easy to see that s_{2k} is trace vanishing.

Example 1.6. $\text{Im } p$ can indeed equal $\{0\}$, K , $\text{sl}_n(K)$, or $M_n(K)$.

- (i) $\text{Im } p = \{0\}$ iff the polynomial p is a PI, and s_{2n} is an example of such a polynomial by the Amitsur–Levitzki theorem [3]. To determine all such PI's leads us to Specht's finite basis problem, which for multilinear polynomial identities is not yet settled in nonzero characteristic. Also Kemer's solution is non-computational, so we do not yet have explicit generators of the T-ideal of PI's for $n > 2$.
- (ii) If the polynomial p is central, then its image is K , and examples of such polynomials can be found in [89] and in [41]. For $n = 2$ the central polynomial of smallest degree is the multilinearization of $[x, y]^2$. The central polynomial of smallest degrees are known for $n = 3, 4$ but not in general.
- (iii) When $p = x_1x_2 - x_2x_1$, $\text{Im } p = \text{sl}_n(K)$ by a theorem of Albert and Muckenhoupt [1]. Obviously p has the lowest possible degree.

In general, if $\text{Im } p \subseteq \text{sl}_n(K)$, then p is trace vanishing, which raises the issue, when does this imply that $\text{Im } p$ consists of all commutators? We shall investigate this issue, and obtain counterexamples for non-multilinear polynomials.

- (iv) $\text{Im } p = M_n(K)$ for $p = x$.

When K is a finite field, Chuang [23] proved that any subset $S \subseteq M_n(K)$ containing 0 is the image of a polynomial with constant term zero, if and only if S is invariant under conjugation. Later Chuang's result was generalized by Kulyamin [68, 69] for graded algebras.

The research detailed in this paper focuses on associative algebras, mostly matrix algebras $M_n(K)$ over an infinite field K , for $n = 2, 3$. We also have density results for arbitrary n . In [36] Dykema and Klep obtain an affirmative answer for the L'vov–Kaplansky conjecture, for multilinear polynomials of degree 3 when n is either even, or odd and ≤ 15 .

In [85] Mesyan conjectured that if $n \geq m - 1$, then any multilinear polynomial of degree m evaluated on $M_n(K)$ takes all values of trace zero, and proved it for $m = 3$. In [22] Buzinski and Winstenley proved Mesyan's conjecture for $m = 4$.

In [39] Fagundes denotes by $UT_n^{(k)}$ for $k \geq 0$ the set of strictly upper triangular matrices which, besides the main diagonal, also have k zero diagonals located above the main diagonal, and proves that if p is a multilinear polynomial evaluated on $UT_n^{(0)}$ of degree m then its image is either $\{0\}$ or $UT_n^{(m-1)}$. In particular, $\text{Im } p$ is a vector space.

In [40] Fagundes and de Mello describe the images of multilinear polynomials of degree ≤ 4 on the upper triangular matrices.

Other works on upper triangular matrices include [96, 104, 105, 106].

Vitas [102] proved for any nonzero multilinear polynomial p , that if A is an algebra with a surjective inner derivation, such as the Weyl algebra, then $\text{Im } p = A$.

A *Lie polynomial* is an element of the free Lie algebra in the alphabet $\{x_i : i \in I\}$, cf. [91, p. 8]. In other words, a Lie polynomial is a sum of Lie monomials $\alpha_j h_j$, where h_j is a Lie word, built inductively: each letter x_i is a Lie word of degree 1, and if h_j, h_k are distinct Lie words of degree d_j and d_k , then $[h_j, h_k]$ is a Lie word of degree $d_j + d_k$. A Lie polynomial p is *multilinear* if each letter appearing in p appears exactly once in each of its Lie monomials.

In [103] Špenko proved the L'vov–Kaplansky conjecture for Lie polynomials p of degree ≤ 4 evaluated on matrix algebra M_n .

In [4] Anzis, Emrich and Valiveti proved the L'vov–Kaplansky conjecture for multilinear Lie polynomials of degree 3 and 4 evaluated on the Lie algebras $\text{su}(n)$ of traceless skew-Hermitian matrices and $\text{so}(n)$ of skew-symmetric matrices.

In [78] Ma and Oliva proved that the image of any multilinear Jordan polynomial of degree 3 evaluated on the Jordan algebras of real and complex symmetric matrices forms a vector space.

In [76] Li and Tsui proved that if R is a central simple algebra of degree n over its center F and $\text{char}(F) = 0$, with $\lambda \in F \setminus \{0, -1\}$, then there exist $a, c \in R$ such that for any element $r \in R$ of reduced trace 0 there is an element $b \in R$ such that $r = [a, [c, b]] + \lambda[c, [a, b]]$.

In [17] Brešar proved that for any unital algebra A over a field F of characteristic 0, if $1 \in [A, A]$ then $[A, A] \subseteq \text{span } f(A)$ for every nonconstant polynomial f . Also he investigated the set $f(A) - f(A)$ of differences of evaluations and proved that for any algebraically closed field F of characteristic 0 and any noncommutative polynomial f , the set $f(A) - f(A)$ on $A = M_n(F)$ with $n \geq 2$ contains all square-zero matrices.

Papers on polynomial maps evaluated on matrix algebras include [49, 107], who investigated maps that preserve zeros of multilinear polynomials.

Research into polynomial image sets has strong connections with the PI-theory. In particular, a polynomial identity is a polynomial whose image set is $\{0\}$, and a central polynomial is a polynomial whose images are central elements (in matrix algebras, central elements are scalar matrices). The methods of working with PI are set out in [19, 26, 31, 32, 34, 42, 43, 48, 50, 54, 63, 64, 75, 87, 88, 95], for more detailed exposition of PI theory and related references see [56]. For combinatorial questions see [12].

In the study of evaluations of polynomials on algebras, the approach associated with the investigation of normal bases of algebras seems significant. In this context, the research into Gröbner–Shirshov bases by the Bokut school (see [13, 14, 15]) is of interest, as is the study of evaluations of polynomials on vertex algebras.

1.1 Evaluations of words

The parallel topic in group theory (the images of words in groups) also has been studied extensively, particularly in recent years. Investigation of the image sets of words in pro- p -groups is related to the investigation of Lie polynomials and helped Zelmanov [108] to prove that the free pro- p -group cannot be embedded in the algebra of $n \times n$ matrices when $p \gg n$. (For $p > 2$, the impossibility of embedding the free pro- p -group into the algebra of 2×2 matrices had been proved by Zubkov [110].) The general problem of nonlinearity of the free pro- p -group is related on the one hand with images of Lie polynomials and words in groups, and on the other hand with problems of Specht type, which is of significant current interest.

Let $w = w(x_1, \dots, x_m)$ be an element of the free group $F_m(X)$, where $X = \{x_1, x_2, \dots, x_m\}$. Given a group G , we consider the corresponding evaluation map $f_{w,G}: G^m \rightarrow G$ corresponding to the word w . This map is called a *word map*, which for convenience we also notate as w instead of $f_{w,G}$. Note that the identity matrix I belongs to the image of any word map.

The major question under consideration is the size of the image $w(G) \subseteq G$. Surjectivity of the map w means that $w(G) = G$, i.e., the equations $w(x_1, \dots, x_m) = g$ can be solved for each $g \in G$. This is, of course, a rare phenomenon even for “good” classes of groups. So usually one has to vary the word w and group G to obtain a reasonable estimate of $w(G)$. The typical classes of groups which provide such estimates are simple algebraic groups, simple and perfect finite groups, and some others.

The theorem of Borel [16] (also cf. [70]) states that for any simple (semisimple) algebraic group G and any word w of the free group on m variables, the word map $w: G^m \rightarrow G$ is dominant. If the ground field K is algebraically closed this implies immediately that $w(G(K))^2 = G$, that is, every element $g \in G$ is a product of two elements from $w(G)$.

For an arbitrary infinite field K and arbitrary word w , Hui–Larsen–Shalev [53] proved that $w(G(K))^4 = G$, that is, every element $g \in G$ is a product of four elements from $w(G)$. This estimate was improved by Egorchenova–Gordeev in [37] to $w(G(K))^3 = G$.

The most challenging open problem for word maps on semisimple algebraic groups is the following, see [65]:

Conjecture 1.7. *Let $G = \mathrm{PSL}_2(\mathbb{C})$, and let $w = w(x, y)$ be an arbitrary non-identity word in $F(x, y)$. Then the word map $w: \mathrm{PSL}_2(\mathbb{C}) \times \mathrm{PSL}_2(\mathbb{C}) \rightarrow \mathrm{PSL}_2(\mathbb{C})$ is surjective. In other words, the equation*

$$w(x_1, x_2) = a$$

has a solution for every $a \in \mathrm{PSL}_2(\mathbb{C})$.

This conjecture is still widely open, there being only several partial results, see [6, 10, 45, 46, 47, 65]. It is a special case of the conjecture in [57, Question 2].

Problem 1.8. *Let \mathbb{G} be the class of simple groups G of the form $G = G(K)$ where $K = \bar{K}$ is an algebraically closed field and G is a semisimple adjoint linear algebraic group. Is it true that word maps evaluated on groups from \mathbb{G} are surjective for all nontrivial non-power words?*

The latter problem for groups of type A_n can be reformulated as follows:

Problem 1.9. *Is the word map $w: \mathrm{PSL}_n(\mathbb{C}) \times \mathrm{PSL}_n(\mathbb{C}) \rightarrow \mathrm{PSL}_n(\mathbb{C})$ surjective for any non-trivial $w(x, y) \in F_2(x, y)$?*

Now we turn from simple algebraic groups to finite simple groups. Ongoing interest to this area was initially stimulated by the positive solution of Ore’s problem: Every element of a finite simple group is a single commutator (see [38] and the final solution in Liebeck–O’Brien–Shalev–Tiep [77]; a survey is given in [84]).

Formidable progress in the description of images of word maps on finite simple groups was obtained by M. Larsen and A. Shalev, who stimulated the development of this area of research under the name “Waring type problems”. The latest result of Larsen–Shalev–Tiep [72] (see also [71, 73, 74]), is as follows:

Theorem 1.10 ([72]). *Let w be an arbitrary non-trivial word of $F(x_1, \dots, x_n)$. There exists a constant $N = N(w)$ such that for all finite non-abelian simple groups of order greater than N one has*

$$w(G)^2 = G.$$

Different aspects of word maps are considered in a vast and extended literature; we refer to the papers [7, 8, 9, 20, 21, 45, 46, 47, 57, 72, 97, 98, 99, 101] for details, surveys and further explanations. Waring type questions for rings were considered by Matei Brešar [17].

In [55] Kanel-Belov, Grigoriev, Elishev and Yu prove the possibility of lifting a symplectomorphism to an automorphism of the power series completion of the Weyl algebra of the corresponding rank. They study the problem of lifting polynomial symplectomorphisms in characteristic zero to automorphisms of the Weyl algebra, by means of approximation by tame automorphisms.

1.2 Non-multilinear polynomials

As noted above, the analog to the L’vov–Kaplansky conjecture formulated for any polynomial fails when K is a finite field, so we may assume that K is infinite. The situation is considerably subtler for images of non-multilinear polynomials.

Definition 1.11. A polynomial p (written as a sum of monomials) is called *semi-homogeneous of weighted degree d* with (integer) *weights* (w_1, \dots, w_m) if for each monomial h of p , taking d_j to be the degree of x_j in h , we have

$$d_1 w_1 + \dots + d_n w_n = d.$$

A semi-homogeneous polynomial with weights $(1, 1, \dots, 1)$ is called *homogeneous* of degree d .

A polynomial p is *completely homogeneous* of multidegree (d_1, \dots, d_m) if each variable x_i appears the same number of times d_i in all monomials. A polynomial $p \in K\langle x_1, \dots, x_m \rangle$ is multilinear iff it is homogeneous of multidegree $(1, 1, \dots, 1)$. Thus, a polynomial is multilinear if it is a polynomial of the form

$$p(x_1, \dots, x_m) = \sum_{\sigma \in S_m} c_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)},$$

where S_m is the symmetric group in m letters and the coefficients c_σ are constants in K .

1.3 The main theorems

Here is a general combinatorial result.

Theorem 1.12 ([60, Theorem 1]). *Let $p(x_1, \dots, x_m)$ be any multilinear polynomial evaluated on $n \times n$ matrices over an infinite field. Assume that p is neither scalar nor PI. Then $\text{Im } p$ contains a matrix of the form $c_n e_{n,1} + \sum_{i=1}^{n-1} c_i e_{i,i+1}$ where $c_1, \dots, c_n \neq 0$. When $\text{char}(K)$ is 0 or prime to n , $\text{Im } p$ contains a matrix with eigenvalues $\{c, c\varepsilon, \dots, c\varepsilon^{n-1}\}$ for some $0 \neq c \in K$, where ε is a primitive n root of 1.*

1.3.1 The main theorems for $n = 2$

Our most decisive results are for $n = 2$, given in [58], for which we settle Conjecture 1.4, proving the following results (see [58, Section 2] for terminology). We call a field K *quadratically closed* (with respect to the polynomial p) if every nonconstant polynomial in $K[x]$ in one variable, of degree $\leq 2 \deg p$, has a root in K .

Theorem 1.13 ([58, Theorem 1]). *Let $p(x_1, \dots, x_m)$ be a semi-homogeneous polynomial evaluated on the algebra $M_2(K)$ of 2×2 matrices over a quadratically closed field. Then $\text{Im } p$ is either $\{0\}$, K , the set of all non-nilpotent matrices having trace zero, $\text{sl}_2(K)$, or a dense subset of $M_2(K)$ (with respect to Zariski topology).*

(We also give examples to show how p can have these images.)

Theorem 1.14 ([58, Theorem 2]). *If p is a multilinear polynomial evaluated on the matrix ring $M_2(K)$ (where K is a quadratically closed field), then $\text{Im } p$ is either $\{0\}$, K , sl_2 , or $M_2(K)$.*

The L'vov–Kaplansky conjecture is proved in [81] for $M_2(\mathbb{R})$, together with a partial solution settling the major part of L'vov–Kaplansky's conjecture in this case, proving the following result:

Theorem 1.15 ([81, Theorem 1]). *If p is a multilinear polynomial evaluated on the matrix ring $M_2(K)$ (where K is an arbitrary field), then $\text{Im } p$ is either $\{0\}$, or K (the set of scalar matrices), or $\text{sl}_2 \subseteq \text{Im } p$. If $K = \mathbb{R}$ then $\text{Im } p$ is either $\{0\}$, or K , or sl_2 or M_2 .*

Remark 1.16. Assume that p is a multilinear polynomial evaluated on 2×2 matrices. According to Theorem 1.15, $\text{Im } p$ is $\{0\}$, or K , or $\text{sl}_2(K)$ or $\text{sl}_2(K) \subsetneq \text{Im } p$. In the last case it is clear that $\text{Im } p$ must be Zariski dense in $M_2(K)$, because otherwise $\dim(\text{Im } p) = 3$ and $\text{Im } p$ is reducible, a contradiction.

The situation is considerably subtler for images of non-multilinear, completely homogeneous polynomials than for multilinear polynomials, but nevertheless a classification of the possible images of homogeneous polynomials evaluated on 2×2 matrices is provided:

Theorem 1.17 ([81, Theorem 2]). *Let $p(x_1, \dots, x_m)$ be a semi-homogeneous polynomial evaluated on 2×2 matrices with real entries. Then $\text{Im } p$ is either:*

- $\{0\}$,
- the set $\mathbb{R}_{\geq 0}$, i.e., the matrices λI for $\lambda \geq 0$,
- the set \mathbb{R} of scalar matrices,
- the set $\mathbb{R}_{\leq 0}$, i.e., the matrices λI for $\lambda \leq 0$,
- the set $\mathrm{sl}_{2,\geq 0}(\mathbb{R})$ of trace zero matrices with non-negative discriminant,
- the set $\mathrm{sl}_{2,\leq 0}(\mathbb{R})$ of trace zero matrices with non-positive discriminant,
- the set $\mathrm{sl}_2(\mathbb{R})$,
- or is Zariski dense in $M_2(\mathbb{R})$.

1.3.2 The main theorems for $n = 3$

We have not classified the possible images of all homogeneous or semi-homogeneous polynomials, but for $n = 3$ we do describe all possible images of trace vanishing semi-homogeneous polynomials:

Theorem 1.18 ([59, Theorem 3]). *Let $p(x_1, \dots, x_m)$ be a semi-homogeneous polynomial which is trace vanishing on 3×3 matrices. Then $\mathrm{Im} p$ is one of the following:*

- $\{0\}$,
- the set of scalar matrices (which can occur only if $\mathrm{Char} K = 3$),
- a dense subset of $\mathrm{sl}_3(K)$, or
- the set of 3-scalar matrices, i.e., the set of matrices having eigenvalues $\{\gamma, \gamma\varepsilon, \gamma\varepsilon^2\}$, where ε is a primitive cube root of 1.

All of the cases in Theorem 1.18 occur, and we give an example of a completely homogeneous 3-scalar polynomial. Unfortunately the question of whether there exists a 3-scalar multilinear polynomial remains open.

Although we do not settle the L'vov–Kaplansky conjecture completely, we describe the possible images of multilinear polynomials.

Theorem 1.19 ([59, Theorem 4]). *Let p be a multilinear polynomial which is trace vanishing on 3×3 matrices over a field K of arbitrary characteristic. Then $\mathrm{Im} p$ is one of the following:*

- $\{0\}$,
- the set of scalar matrices,
- the set of 3-scalar matrices, or
- for each triple $\lambda_1 + \lambda_2 + \lambda_3 = 0$ there exist a matrix $M \in \mathrm{Im} p$ with eigenvalues λ_1, λ_2 and λ_3 .

Theorem 1.20 ([59, Theorem 2]). *If p is a multilinear polynomial evaluated on 3×3 matrices over algebraically closed field K , then $\mathrm{Im} p$ is one of the following:*

- $\{0\}$,
- the set of scalar matrices,
- $\mathrm{sl}_3(K)$, (perhaps lacking the diagonalizable matrices of discriminant 0),
- a dense subset of $M_3(K)$,
- the set of 3-scalar matrices, or
- the set of sums of scalar and 3-scalar matrices.

1.3.3 Higher n

Let us improve the estimates of the dimension of $\text{Im } p$, first for $n \geq 5$ and then for $n = 4$.

Theorem 1.21 ([60, Theorem 4]). *Assume that the characteristic of K does not divide n , and $K = F[\varepsilon]$, where ε is a primitive n -th root of 1. Let p be any multilinear polynomial evaluated on $n \times n$ matrices which is not PI or central. If $n \geq 5$, then the image of p is at least $(n^2 - n + 3)$ -dimensional.*

Theorem 1.22 ([60, Theorem 5]). *Let p be any multilinear polynomial evaluated on 4×4 matrices, which is neither PI nor central. Assume that $\text{Char } K \neq 2$. Then $\dim \text{Im } p \geq 14$, equality holding only if the following conditions are satisfied:*

- *For any matrix units a_i , if $p(a_1, \dots, a_m)$ is diagonal then it has eigenvalues $(c, c, -c, -c)$ for some c .*
- *Any value of p has eigenvalues $(\lambda_1, \lambda_2, -\lambda_1, -\lambda_2)$.*

In [61] we investigated possible images of homogeneous Lie polynomials evaluated on 2×2 matrices $n = 2$, and obtained the following result:

Theorem 1.23 ([61, Theorem 3]). *For any algebraically closed field K of characteristic $\neq 2$, the image of any Lie polynomial f (not necessarily homogeneous) evaluated on $\text{sl}_2(K)$ is either $\text{sl}_2(K)$, or $\{0\}$, or the set of trace zero non-nilpotent matrices.*

We also provide examples, showing the existence of completely homogeneous Lie polynomials with exactly these image sets. In [61] adjoint maps are used to construct an important example of the completely homogeneous Lie polynomial which image is the set of all trace vanishing matrices except for the nilpotents.

1.3.4 Polynomials over quaternion algebras

One can generalize the L'vov–Kaplansky conjecture for an arbitrary finite dimensional simple algebra.

Definition 1.24. By the quaternion algebra \mathbb{H} we mean the four-dimensional algebra $\langle 1, i, j, k \rangle_{\mathbb{R}}$ such that

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j.$$

In this algebra, $\mathbb{R} = \mathbb{R}1$ are called *scalars*, and $V = \mathbb{R}i + \mathbb{R}j + \mathbb{R}k$ are called *pure quaternions*; $\{1, i, j, k\}$ are called *basic quaternions*. We also will use the standard quaternion functions: the *norm* $\|a + bi + cj + dk\| = \sqrt{a^2 + b^2 + c^2 + d^2}$, the *real part* $\text{Re}(a + bi + cj + dk) = a$, and the *pure quaternion part* $\text{Ve}(a + bi + cj + dk) = bi + cj + dk$.

Any quaternion can be uniquely written as a sum of a scalar and a pure quaternion, hence the functions of real and pure quaternion parts are well defined. The norm function is multiplicative.

In [2] Almutairi proved that if p is a non-central multilinear polynomial then its image contains all pure quaternions. In Section 2.1.5 we provide a complete classification of $\text{Im } p$, settling the L'vov–Kaplansky conjecture for the quaternion algebra

Theorem 1.25 ([82, Theorem 1]). *If p is a multilinear polynomial evaluated on the quaternion algebra $\mathbb{H}(\mathbb{R})$, then $\text{Im } p$ is either $\{0\}$, \mathbb{R} (the space of scalar quaternions), or V (the space of pure quaternions), or $\mathbb{H}(\mathbb{R})$.*

Also, the matrix representation of the quaternions ring is used in order to provide a classification of possible images of semi-homogeneous polynomials evaluated on the quaternion algebra.

Theorem 1.26 ([82, Theorem 2]). *If p is a semi-homogeneous polynomial evaluated on the quaternion algebra $\mathbb{H}(\mathbb{R})$, then $\text{Im } p$ is either $\{0\}$, or \mathbb{R} , or $\mathbb{R}_{\geq 0}$, or $\mathbb{R}_{\leq 0}$, or V , or some Zariski dense subset of \mathbb{H} .*

1.4 Some basic tools

We recall the following elementary graph-theoretic lemma.

Lemma 1.27. *Let p be a multilinear polynomial. If the a_i are matrix units, then $p(a_1, \dots, a_m)$ is either 0, or $c \cdot e_{ij}$ for some $i \neq j$, or a diagonal matrix.*

The proof of this lemma is presented in Section 3.4.

Lemma 1.28. *If $\text{Im } p$ consists only of diagonal matrices, then $\text{Im } p$ is either $\{0\}$ or the set K of scalar matrices.*

Proof. Suppose that some nonscalar diagonal matrix $A = \text{Diag}\{\lambda_1, \dots, \lambda_n\}$ is in the image. Therefore $\lambda_i \neq \lambda_j$ for some i and j . The matrix $A' = A + e_{ij}$ (here e_{ij} is the matrix unit) is conjugate to A so by Remark 1.1 also belongs to $\text{Im } p$. However A' is not diagonal, a contradiction. ■

The proofs of our theorems use some algebraic-geometric tools in conjunction with these ideas from graph theory. The final part of the proofs of Theorems 1.14 and 1.20 uses the First fundamental theorem of invariant theory (that in case $\text{Char } K = 0$, invariant functions evaluated on matrices are polynomials involving traces), proved by Helling [51], Procesi [86], and Razmyslov [90]. The formulation in positive characteristic, due to Donkin [30], is somewhat more intricate. The group $\text{GL}_n(K)$ acts on m -tuples of $n \times n$ -matrices by simultaneous conjugation.

Theorem (Donkin [30]). *For any $m, n \in \mathbb{N}$, the algebra of polynomial invariants*

$$K[M_n(K)^m]^{\text{GL}_n(K)}$$

under $\text{GL}_n(K)$ is generated by the trace functions

$$T_{i,j}(x_1, x_2, \dots, x_m) = \text{Trace} \left(x_{i_1} x_{i_2} \cdots x_{i_r}, \bigwedge^j K^n \right), \quad (1.1)$$

where $i = (i_1, \dots, i_r)$, all $i_l \leq m$, $r \in \mathbb{N}$, $j > 0$, and $x_{i_1} x_{i_2} \cdots x_{i_r}$ acts as a linear transformation on the exterior algebra $\bigwedge^j K^n$.

Remark 1.29. For $n = 2$, we have a polynomial function in expressions of the form

$$\text{Trace} \left(A, \bigwedge^2 K^2 \right)$$

and $\text{tr } A$ where A is monomial. ($\text{tr } A$ denotes the trace.) Note that

$$\text{Trace} \left(A, \bigwedge^2 K^2 \right) = \det A.$$

We also need the first fundamental theorem of invariant theory (see [86, Theorem 1.3]).

Proposition 1.30. *Any polynomial invariant of $n \times n$ matrices A_1, \dots, A_m is a polynomial in the invariants $\text{tr}(A_{i_1} A_{i_2} \cdots A_{i_k})$, taken over all possible (noncommutative) products of the A_i .*

(The second fundamental theorem, dealing with relations between invariants, was proved by Procesi [86] and Razmyslov [90] in the case $\text{Char } K = 0$ and by Zubkov [110] in the case $\text{Char } K > 0$. We recommend reader to read the book [27].)

1.4.1 Generic matrices

Another major tool is Amitsur's theorem [92, Theorem 3.2.6, p. 176]. *Generic matrices* over K are $n \times n$ matrices whose entries are commuting indeterminates over K . By Proposition 1.31) the algebra of generic matrices is a domain UD, whose ring of fractions is a division algebra of degree n ,

$$\widetilde{\text{UD}} \subseteq M_n(F(\xi_{i,j}^{(k)}): 1 \leq i, j \leq n, k \geq 1).$$

In fact, we need a slight modification of this theorem, which is well known. We can define the reduced characteristic coefficients of elements of $\widetilde{\text{UD}}$, which by [94, Remark 24.67] lie in F_1 .

Proposition 1.31. *The algebra of generic matrices with traces is a domain which can be embedded in the division algebra $\widetilde{\text{UD}}$ of central fractions of Amitsur's algebra of generic matrices. Likewise, all of the functions in Donkin's theorem can be embedded in $\widetilde{\text{UD}}$.*

Proof. Any trace function can be expressed as the ratio of two central polynomials, in view of [92, Theorem 1.4.12]; also see [62, Theorem J, p. 27] which says for any characteristic coefficient ω_k of the characteristic polynomial

$$\lambda^t + \sum_{k=1}^t (-1)^k \omega_k \lambda^{t-k}$$

that

$$\omega_k f(a_1, \dots, a_t, r_1, \dots, r_m) = \sum_{k=1}^t f(T^{k_1} a_1, \dots, T^{k_t} a_t, r_1, \dots, r_m), \quad (1.2)$$

summed over all vectors (k_1, \dots, k_t) where each $k_i \in \{0, 1\}$ and $\sum k_i = k$, where f is any t -alternating polynomial (and $t = n^2$). In particular, taking $k = 1$, so that $\omega_k = \text{tr}(T)$, we have

$$\text{tr}(T) f(a_1, \dots, a_t, r_1, \dots, r_m) = \sum_{j=1}^t f(a_1, \dots, a_{j-1}, T a_j, a_{j+1}, \dots, a_t, r_1, \dots, r_m),$$

so the trace

$$\text{tr}(T) = \frac{\sum_{j=1}^t f(a_1, \dots, a_{j-1}, T a_j, a_{j+1}, \dots, a_t, r_1, \dots, r_m)}{f(a_1, \dots, a_t, r_1, \dots, r_m)}$$

belongs to $\widetilde{\text{UD}}$. In general, the function (1.1) of Donkin's theorem is a matrix invariant and thus can be written in terms of characteristic coefficients of matrices, so we can apply equation (1.2). ■

Lemma 1.32. *If $\text{Char } K$ does not divide n , then any non-identity $p(x_1, \dots, x_m)$ of $M_n(K)$ must either be a central polynomial or take on a value which is a matrix whose eigenvalues are not all the same.*

Proof. Otherwise $p(x_1, \dots, x_m) - \frac{1}{n} \text{tr}(p(x_1, \dots, x_m))$ is a nilpotent element in the algebra of generic matrices with traces, so by Proposition 1.31 is 0, implying p is central. ■

For $n > 2$, we have an easy consequence of the theory of division algebras.

Lemma 1.33. *Suppose for some polynomial p and some number $q < n$, that p^q takes on only scalar values in $M_n(K)$, over an infinite field K , for n prime. Then p takes on only scalar values in $M_n(K)$.*

Proof. We can view p as an element of the generic division algebra $\widetilde{\text{UD}}$, and we adjoin a q -root of 1 to K if necessary. Then p generates a subfield of dimension 1 or n of $\widetilde{\text{UD}}$. The latter is impossible, so the dimension is 1; i.e., p is already central. ■

We also require one basic fact from linear algebra:

Lemma 1.34. *Let V_i (for $1 \leq i \leq m$) and V be linear spaces over arbitrary field K . Let $f(T_1, \dots, T_m): \prod_{i=1}^m V_i \rightarrow V$ be a multilinear mapping (i.e., linear with respect to each T_i). Assume there exist two points in $\text{Im } f$ which are not proportional. Then $\text{Im } f$ contains a 2-dimensional plane. In particular, if V is 2-dimensional, then $\text{Im } f = V$.*

Proof. Let us denote for $\mu = (T_1, \dots, T_m)$ and $\nu = (T'_1, \dots, T'_m) \in \prod_{i=1}^m V_i$

$$\text{Dist}(\mu, \nu) = \#\{i: T_i \neq T'_i\}.$$

Consider $k = \min \{d: \text{there exists } \mu, \nu \in \prod_{i=1}^m V_i \text{ such that } f(\mu) \text{ is not proportional to } f(\nu) \text{ and } \text{Dist}(\mu, \nu) = d\}$. We know $k \leq m$ by assumptions of lemma. Also $k \geq 1$ since any element of V is proportional to itself. Assume $k = 1$. In this case there exist i and T_1, \dots, T_m, T'_i such that $f(T_1, \dots, T_m)$ is not proportional to $f(T_1, \dots, T_{i-1}, T'_i, T_{i+1}, \dots, T_m)$. Therefore

$$\langle f(T_1, \dots, T_m), f(T_1, \dots, T_{i-1}, T'_i, T_{i+1}, \dots, T_m) \rangle \subseteq \text{Im } p$$

is 2-dimensional. Hence we can assume $k \geq 2$. We can enumerate variables and consider $\mu = (T_1, \dots, T_m)$ and $\nu = (T'_1, \dots, T'_k, T_{k+1}, \dots, T_m)$, $v_1 = f(\mu)$ is not proportional to $v_2 = f(\nu)$. Take any $a, b \in K$. Consider $v_{a,b} = f(aT_1 + bT'_1, T_2 + T'_2, \dots, T_k + T'_k, T_{k+1}, \dots, T_m)$. Let us open the brackets. We have

$$v_{a,b} = av_1 + bv_2 + \sum_{\emptyset \subsetneq S \subsetneq \{1, \dots, k\}} c_S f(\theta_S),$$

where c_S equals a if $1 \in S$ and b otherwise, and $\theta_S = (\tilde{T}_1, \dots, \tilde{T}_k, T_{k+1}, \dots, T_m)$ for $\tilde{T}_i = T_i$ if $i \in S$ or T'_i otherwise. Note that any θ_S in the sum satisfies $\text{Dist}(\theta_S, \mu) < k$ and $\text{Dist}(\theta_S, \nu) < k$ therefore $f(\theta_S)$ must be proportional to both v_1 and v_2 and thus $f(\theta_S) = 0$. Therefore $v_{a,b} = av_1 + bv_2$ and hence $\text{Im } f$ contains a 2-dimensional plane. ■

Lemma 1.35. *The multiplicity of any eigenvalue of an element a of $\widetilde{\text{UD}}$ must divide n . In particular, when n is odd, a cannot have an eigenvalue of multiplicity 2.*

Proof. Recall [93, Remark 4.106] that for any element a in a division algebra, represented as a matrix, the eigenvalues of a occur with the same multiplicity, which thus must divide n . ■

Lemma 1.36. *Assume that an element a of $\widetilde{\text{UD}}$ has a unique eigenvalue α (of multiplicity n). If $\text{Char}(K) = 0$, then a is scalar.*

If $\text{Char}(K) = k \neq 0$, then a is k^l -scalar for some l .

Proof. If $\text{Char}(K) = 0$, then α is an element of $\widetilde{\text{UD}}$ and $a - \alpha I$ is nilpotent, and thus 0.

If $\text{Char}(K) = k$ then α^{k^l} is an element of $\widetilde{\text{UD}}$, therefore $a^{k^l} - \alpha^{k^l} I$ is nilpotent, and thus 0. Thus a is k^l -scalar. This is impossible unless k divides n . ■

Over any field K , applying the structure theory of division rings to Amitsur's theorem, it is not difficult to get an example of a completely homogeneous polynomial f , noncentral on $M_3(K)$, whose values all have third powers central; clearly its image does not comprise a subspace of $M_3(K)$. Furthermore, in the (non-multilinear) completely homogeneous case, the set of values could be dense without including all matrices. (Analogously, although the finite basis problem for multilinear identities is not yet settled in nonzero characteristic, there are counterexamples for completely homogeneous polynomials, cf. [11].)

1.4.2 Generic elements

Definition 1.37. Assume that K is an arbitrary field and $F \subseteq K$ is a subfield. The set $\{\xi_1, \dots, \xi_k\} \subseteq K$ is called *generic* (over F) if $f(\xi_1, \dots, \xi_k) \neq 0$ for any commutative polynomial $f \in F[x_1, \dots, x_k]$ that takes nonzero values.

Lemma 1.38. Assume that K has infinite transcendence degree over F . Then for any $k \in \mathbb{N}$ there exists a set of generic elements $\{\xi_1, \dots, \xi_k\} \subset K$.

Proof. K has infinite transcendence degree over F . Therefore, there exists an element $\xi_1 \in K \setminus \bar{F}$, where \bar{F} is an algebraic closure of F . Now we consider $F_1 = F[\xi_1]$. K has infinite transcendence degree over F and thus has infinite transcendence degree over F_1 . Therefore there exists an element $\xi_2 \in K \setminus \bar{F}_1$. And we consider the new base field $F_2 = F_1[\xi_2]$. We can continue up to any natural number k . ■

Remark 1.39. Note according to Lemma 1.38 that if K has infinite transcendence degree over F we can take as many generic elements as we need. In particular we can take as many generic matrices as we need.

Lemma 1.40. Assume $f: H \rightarrow \mathbb{R}$ (where $H \subseteq \mathbb{R}^k$ is an open set in k -dimensional Euclidean space) is a function that is continuous in a neighborhood of the point $(y_1, \dots, y_k) \in H$, with $f(y_1, \dots, y_k) < q$. Let c_i be real numbers (in particular the coefficients of some polynomial p). Then there exists a set of elements $\{x_1, \dots, x_k\} \subseteq \mathbb{R}$ generic over $F = \mathbb{Q}[c_1, \dots, c_N]$ such that $(x_1, \dots, x_k) \in H$ and $f(x_1, \dots, x_k) < q$.

Proof. The δ -neighborhood $N_\delta(x)$ of $x \in \mathbb{R}$ denotes the interval $(x - \delta, x + \delta) \subseteq \mathbb{R}$. Fix some small $\delta > 0$ such that the product of δ -neighborhoods of y_k lays in H . For this particular δ we consider the δ -neighborhood $N_\delta(y_1)$ of y_1 : the interval $(y_1 - \delta, y_1 + \delta)$ is an uncountable set, and therefore there exists $x_1 \in N_\delta(y_1) \setminus \bar{F}$. We consider $F_1 = F[x_1]$ and analogically chose $x_2 \in N_\delta(y_2) \setminus \bar{F}_1$ and take $F_2 = F_1[x_2]$. In such a way we can take generic elements $x_k \in N_\delta(y_k)$. Note that if δ is not sufficiently small $f(x_1, \dots, x_k)$ can be larger than q , but

$$f(x_1, \dots, x_k) \xrightarrow{\delta \rightarrow 0} f(y_1, \dots, y_k).$$

Thus there exists sufficiently small δ and generic elements $x_i \in N_\delta(y_i)$ such that $f(x_1, \dots, x_k) < q$. ■

Remark 1.41. Note that f can be a function defined on a set of matrices. In this case we consider it as a function defined on the matrix entries.

Remark 1.42. Assume that $\text{Char}(K) = 0$. Suppose t is a commuting indeterminate, and $f(x_1, \dots, x_m; t)$ is a polynomial taking values under matrix substitutions for the x_i and scalars for t . If there exists unique t_0 such that $f(x_1, \dots, x_m; t_0) = 0$, then t_0 is a rational function with respect to the entries of x_i . If this t_0 is fixed under simultaneous conjugation of the matrices x_1, \dots, x_m , then t_0 is in the center of Amitsur's generic division algebra $\widehat{\text{UD}}$, implying $f \in \widehat{\text{UD}}$. If $\text{Char}(K) = k \neq 0$, then $t_0^{k^l}$ is a rational function for some $l \in \mathbb{N}_0$. For details see [58, Remark 2].

Remark 1.43. In Remark 1.42 we could take a system of polynomial equations and polynomial inequalities. If t_0 is unique, then it is a rational function (or t_0^k if $\text{Char}(K) = k$).

1.4.3 Basic facts about cones

Here is one of the main tools for our investigation.

Definition 1.44. A *cone* of $M_n(K)$ is a subset closed under multiplication by nonzero constants. An *invariant cone* is a cone invariant under conjugation. An invariant cone is *irreducible* if it does not contain any nonempty invariant cone.

Example 1.45. Examples of invariant cones of $M_n(K)$ include:

- (i) The set of diagonalizable matrices.
- (ii) The set of non-diagonalizable matrices.
- (iii) The set K of scalar matrices.
- (iv) The set of nilpotent matrices.
- (v) The set sl_n of matrices having trace zero.

Let us continue with the following easy but crucial lemma.

Lemma 1.46. *Suppose the field K is closed under d -roots. If the image of a semi-homogeneous polynomial p of weighted degree d intersects an irreducible invariant cone C nontrivially, then $C \subseteq \text{Im } p$.*

Proof. If $A \in \text{Im } p$ then $A = p(x_1, \dots, x_m)$ for some $x_i \in M_n(K)$. Thus for any $c \in K$, $cA = p(c^{w_1/d}x_1, c^{w_2/d}x_2, \dots, c^{w_m/d}x_m) \in \text{Im } p$, where (w_1, \dots, w_m) are the weights. This shows that $\text{Im } p$ is a cone. ■

Remark 1.47. When the polynomial p is multilinear, the image of any multilinear polynomial is an invariant cone, without any assumption on K .

1.5 Some general observations

Lemma 1.48. *Let K be an algebraically closed field of characteristic 0, and let I be an ideal of $K[X_1, \dots, X_n]$, and $V(I) = \{x \in K^n : f(x) = 0 \forall f \in I\}$. Let $\pi: K^n \rightarrow K^{n-1}$ be the projection onto the first $n-1$ coordinates. Let I' denote the ideal $I \cap K[X_1, \dots, X_{n-1}]$ of $K[X_1, \dots, X_{n-1}]$. Then:*

- (1) $\pi(V(I))$ is a Zariski dense subset of $V(I')$;
- (2) if there exists a Zariski dense subset W of $V(I')$ such that the pre-image $\pi^{-1}(p) \cap V(I)$ of each point $p \in W$ consists of one point, then there exists a rational K -valued function ϕ on $V(I')$ such that all points of a Zariski-dense subset of $V(I)$ have the form $(p, \phi(p))$ where $p \in V(I')$.

If $\text{Char}(K) = k > 0$ then there exists non-negative integer ℓ such that $(p, a) \in V(I)$ satisfy $\phi(p) = a^{k^\ell}$ on a Zariski-dense subset of $V(I)$.

Proof. (1) is by [25, Chapter 3, Section 2, Theorem 3] and the subsequent remarks.

To prove (2), note that by (1) π induces a field homomorphism (hence an embedding) $K(V(I')) \rightarrow K(V(I))$ between the fields of rational functions on the respective varieties. It is enough to show that this is an isomorphism. Indeed, $K(V(I))$ is generated by $K(V(I'))$ and X_n . Moreover, X_n is algebraic over $K(V(I'))$. Let h be the minimal polynomial of X_n

over $K(V(I'))$, of degree d . The derivative h' has degree $d - 1$, and the discriminant $\text{Discr}(h)$ is, up to a scalar, the resultant of h and h' ; it is non-zero since h is irreducible implies that h and h' are relatively prime. Let now U be the open subset of $V(I')$ in which $\text{Discr}(h) \neq 0$ and the coefficients of h are defined. Then each point of U has precisely d distinct π -preimages in $V(I)$. It follows that $d = 1$, as required.

If $\text{Char}(K) = k > 0$, we take ℓ such that $h(x) = h_1(x^{k^\ell})$ but h'_1 is not identically zero. ■

Lemma 1.49. *Assume that $\text{Char}(K) = 0$. If an element a of $\widetilde{\text{UD}}$ has a unique eigenvalue α (i.e., of multiplicity n), then a is scalar. If $\text{Char}(K) = k \neq 0$ then a is k^l -scalar for some l .*

Proof. If $\text{Char}(K) = 0$, then α is an element of $\widetilde{\text{UD}}$ and $a - \alpha I$ is nilpotent, and thus 0.

If $\text{Char}(K) = k$ then α^{k^l} is an element of $\widetilde{\text{UD}}$, therefore $a^{k^l} - \alpha^{k^l} I$ is nilpotent, and thus 0. Thus a is k^l -scalar. This is impossible if k is not the divisor of the size of the matrices n . ■

Remark 1.50. The variety of $n \times n$ matrices with a given set of n distinct eigenvalues has dimension $n^2 - n$.

Remark 1.51. Assume for some matrix units a_i that $p(a_1, \dots, a_m)$ is a diagonal matrix. Then f as constructed in (4.1) in the proof of Theorem 1.12 is diagonal. If the dimension of $\text{Im } f$ is δ , then each evaluation M of f has some set of eigenvalues, and (if the point is generic and the eigenvalues are distinct), then any matrix with this set of eigenvalues is similar to M and therefore belongs to $\text{Im } p$. Therefore by Remark 1.50, $\text{Im } p$ has dimension at least $n^2 - n + \delta$.

Other works on polynomial maps evaluated on matrix algebras include [49, 107], who investigated maps that preserve zeros of multilinear polynomials.

2 The low rank case

Let p be a multilinear polynomial in several non-commuting variables with coefficients in a quadratically closed field K of arbitrary characteristic. In this section we prove the L'vov–Kaplansky conjecture for $n = 2$ in several general cases, and show that although the analogous assertion fails for completely homogeneous polynomials, one can salvage most of the conjecture by including the set of all non-nilpotent matrices of trace zero and also permitting dense subsets of $M_n(K)$.

Remark 2.1. For $n = 2$, Donkin's theorem provides a polynomial function in expressions of the form $\text{Trace}(A, \wedge^2 K^2)$ and $\text{tr } A$ where A is monomial. Note that $\text{Trace}(A, \wedge^2 K^2) = \det A$.

Next we introduce the cones of main interest to us, drawing from Example 1.45.

Example 2.2.

- (i) The set of nonzero nilpotent matrices comprise an irreducible invariant cone, since these all have the same minimal and characteristic polynomial x^2 .
- (ii) The set of nonzero scalar matrices is an irreducible invariant cone.
- (iii) \tilde{K} denotes the set of non-nilpotent, non-diagonalizable matrices in $M_2(K)$. Note that $A \in \tilde{K}$ precisely when A is non-scalar, but with equal nonzero eigenvalues, which is the case if and only if A is the sum of a nonzero scalar matrix with a nonzero nilpotent matrix. These are all conjugate when the scalar part is the identity, i.e., for matrices of the form

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, \quad a \neq 0$$

since these all have the same minimal and characteristic polynomials, namely $x^2 - 2x + 1$. It follows that \tilde{K} is an irreducible invariant cone.

(iv) \hat{K} denotes the set of non-nilpotent matrices in $M_2(K)$ that have trace zero.

When $\text{Char } K \neq 2$, \hat{K} is an irreducible invariant cone, since any such matrix has distinct eigenvalues and thus is conjugate to $\begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}$.

When $\text{Char } K = 2$, \hat{K} is an irreducible invariant cone, since any such matrix is conjugate to $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

(v) $\text{sl}_2(K) \setminus \{0\}$ is the union of the two irreducible invariant cones of (i) and (iv). (The cases $\text{Char } K \neq 2$ and $\text{Char } K = 2$ are treated separately.)

(vi) Let C denote the set of nonzero matrices which are the sum of a scalar and a nilpotent matrix. Then C is the union of the following three irreducible invariant cones: The nonzero scalar matrices, the nilpotent matrices, and the nonzero scalar multiples of non-identity unipotent matrices. (All non-identity unipotent matrices are conjugate.)

2.1 2×2 matrices

2.1.1 The case of a quadratically closed field

In this subsection we assume that K is a quadratically closed field. In particular, all of the eigenvalues of a matrix $A \in M_2(K)$ lie in K . We start with the semi-homogeneous case (which includes the completely homogeneous case), and finally give the complete picture for the multilinear case.

Theorem 2.3. *Let $p(x_1, \dots, x_m)$ be a semi-homogeneous polynomial evaluated on the algebra $M_2(K)$ of 2×2 matrices over a quadratically closed field. Then $\text{Im } p$ is either $\{0\}$, K , the set of all non-nilpotent matrices having trace zero, $\text{sl}_2(K)$, or a dense subset of $M_2(K)$ (with respect to Zariski topology).*

(We also give examples to show how p can have these images.)

Theorem 2.4. *If p is a multilinear polynomial evaluated on the matrix ring $M_2(K)$ (where K is a quadratically closed field), then $\text{Im } p$ is either $\{0\}$, K , sl_2 , or $M_2(K)$.*

Whereas one has a decisive answer for multilinear polynomials, the situation is ambiguous for homogeneous polynomials, since, as we shall see, certain invariant sets cannot occur as their images. For the general non-homogeneous case, the image of a polynomial need not be dense, even if it is non-central and takes on values of nonzero trace, as we see in Example 2.17.

One of our main ideas is to consider some invariant of the matrices in $\text{Im}(p)$, and study the corresponding invariant cones. Here is the first such invariant that we consider.

Remark 2.5. Any non-nilpotent 2×2 matrix A over a quadratically closed field has two eigenvalues λ_1 and λ_2 , such that at least one of which is nonzero. Therefore one can define the ratio of eigenvalues, which is well-defined up to taking reciprocals: $\frac{\lambda_1}{\lambda_2}$ and $\frac{\lambda_2}{\lambda_1}$. Thus, we will say that two non-nilpotent matrices have *different ratios* of eigenvalues if and only if their ratios of eigenvalues are not equal nor reciprocal.

We do have a well-defined mapping $\Pi: M_2(K) \rightarrow K$ given by $A \mapsto \frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1}$. This mapping is algebraic because

$$\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1} = -2 + \frac{(\text{tr } A)^2}{\det A}.$$

Remark 2.6. The set of non-scalar diagonalizable matrices with a fixed nonzero ratio r of eigenvalues (up to taking reciprocals) is an irreducible invariant cone. Indeed, this is true since any such diagonalizable matrix is conjugate to

$$\lambda \begin{pmatrix} 1 & 0 \\ 0 & r \end{pmatrix}.$$

Images of semi-homogeneous polynomials.

Lemma 2.7. *Suppose K is closed under d -roots, as well as being quadratically closed. If the image $\text{Im } p$ of a semi-homogeneous polynomial p of weighted degree d contains an element of \tilde{K} , then $\text{Im } p$ contains all of \tilde{K} .*

Proof. This is clear from Lemma 1.46 together with Example 2.2(iii), since \tilde{K} is an irreducible invariant cone. ■

For the proof of Theorem 2.3 see Section 2.2.

We illuminate this result with some examples to show that certain cones are obtained from specific completely homogeneous polynomials.

Example 2.8.

- (i) The polynomial $g(x_1, x_2) = [x_1, x_2]^2$ has the property that $g(A, B) = 0$ whenever A is scalar, but g can take on a nonzero value whenever A is non-scalar. Thus, $g(x_1, x_2)x_1$ takes on all values except scalars. This polynomial is completely homogeneous, but not multilinear. (One can linearize in x_2 to make g linear in each variable except x_1 , and the same idea can be applied to Formanek's construction [41] of a central polynomial for any n .)
- (ii) Let S be any finite subset of K . There exists a completely homogeneous polynomial p such that $\text{Im } p$ is the set of all 2×2 matrices except the matrices with ratio of eigenvalues from S . The construction is as follows. Consider

$$f(x) = x \cdot \prod_{\delta \in S} (\lambda_1 - \lambda_2 \delta)(\lambda_2 - \lambda_1 \delta),$$

where $\lambda_{1,2}$ are eigenvalues of x . For each δ the product $(\lambda_1 - \lambda_2 \delta)(\lambda_2 - \lambda_1 \delta)$ is a polynomial of $\text{tr } x$ and $\text{tr } x^2$. Thus $f(x)$ is a polynomial with traces, and, by [92, Theorem 1.4.12]), one can rewrite each trace in f as a fraction of multilinear central polynomials. After that we multiply the expression by the product of all the denominators, which we can take to have value 1. We obtain a completely homogeneous polynomial p which image is the cone under $\text{Im } f$ and thus equals $\text{Im } f$. The image of p is the set of all non-nilpotent matrices with ratios of eigenvalues not belonging to S .

- (iii) The image of a completely homogeneous polynomial evaluated on 2×2 matrices can also be \tilde{K} . Take $f(x, y) = [x, y]^3$. This is the product of $[x, y]^2$ and $[x, y]$. $[x, y]^2$ is a central polynomial, and therefore $\text{tr } f = 0$. However, there are no nonzero nilpotent matrices in $\text{Im } p$ because if $[A, B]^3$ is nilpotent then $[A, B]$ (which is a scalar multiple of $[A, B]^3$) is nilpotent and therefore $[A, B]^2 = 0$ and $[A, B]^3 = 0$.
- (iv) Consider the polynomial

$$p(x_1, x_2, y_1, y_2) = [(x_1 x_2)^2, (y_1 y_2)^2]^2 + [(x_1 x_2)^2, (y_1 y_2)^2][x_1 y_1, x_2 y_2]^2.$$

Then p takes on all scalar values (since it becomes central by specializing $x_1 \mapsto x_2$ and $y_1 \mapsto y_2$), but also takes on all nilpotent values, since specializing $x_1 \mapsto I + e_{12}$, $x_2 \mapsto e_{22}$, and $y_1 \mapsto e_{12}$, and $y_2 \mapsto e_{21}$ sends p to

$$[(e_{12} + e_{22})^2, e_{11}^2]^2 + [(e_{12} + e_{22})^2, e_{11}^2][e_{12}, e_{21}] = 0 - e_{12}(e_{11} - e_{22}) = e_{12}.$$

We claim that $\text{Im } p$ does not contain any matrix $a = p(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2)$ in \tilde{K} . Otherwise, the matrix $[(\bar{x}_1 \bar{x}_2)^2, (\bar{y}_1 \bar{y}_2)^2][\bar{x}_1 \bar{y}_1, \bar{x}_2 \bar{y}_2]^2$ would be the difference of a matrix having equal

eigenvalues and a scalar matrix, but of trace 0, and so would have both eigenvalues 0 and thus be nilpotent. Thus $[(\bar{x}_1\bar{x}_2)^2, (\bar{y}_1\bar{y}_2)^2]$ would also be nilpotent, implying the scalar term $[(\bar{x}_1\bar{x}_2)^2, (\bar{y}_1\bar{y}_2)^2]^2$ equals zero, implying a is nilpotent, a contradiction.

$\text{Im } p$ also contains all matrices having two distinct eigenvalues. We conclude that $\text{Im } p = M_2(K) \setminus \tilde{K}$.

Remark 2.9. In Example 2.8(iv), The intersection S of $\text{Im } p$ with the discriminant surface is defined by the polynomial

$$\text{tr}(p(x_1, \dots, x_m))^2 - 4 \det(p(x_1, \dots, x_m)) = (\lambda_1 - \lambda_2)^2.$$

S is the union of two irreducible varieties (its scalar matrices and the nonzero nilpotent matrices), and thus S is a reducible variety. Thus, we see that the discriminant surface of a polynomial p of the algebra of generic matrices can be reducible, even if it is not divisible by any trace polynomial. Such an example could not exist for p multilinear, since then, by the same sort of argument as given in the proof of Theorem 2.3, the discriminant surface would give a generic zero divisor in Amitsur's generic division algebra $\widetilde{\text{UD}}$ of Proposition 1.31, a contradiction. In fact, we will also see that the image of a multilinear polynomial cannot be as in Example 2.8(iv).

Results for arbitrary polynomials.

Lemma 2.10. *If $A, B \in \text{Im } p$ have different ratios of eigenvalues, then $\text{Im } p$ contains matrices having arbitrary ratios of eigenvalues $\frac{\lambda_1}{\lambda_2} \in K$.*

Proof. If $A = p(x_1, \dots, x_m)$, $B = p(y_1, \dots, y_m) \in \text{Im } p$ have different ratios of eigenvalues, then we can lift the $x_1, \dots, x_m, y_1, \dots, y_m$ to generic matrices, and then $p(x_1, \dots, x_m) = \tilde{A}$ and $p(y_1, \dots, y_m) = \tilde{B}$ also have different ratios of eigenvalues. Then take

$$f(T_1, T_2, \dots, T_m) = p(\tau_1 x_1 + t_1 y_1, \dots, \tau_m x_m + t_m y_m),$$

where $T_i = (t_i, \tau_i) \in K^2$. The polynomial f is linear with respect to all T_i .

In view of Remark 2.9, it is enough to show that the ratio $\frac{(\text{tr } f)^2}{\det f}$ takes on all values. Fix $T_1, \dots, T_{i-1}, T_{i+1}, \dots, T_m$ to be generic pairs where i is such that $\frac{(\text{tr } f)^2}{\det f}$ is not constant with respect to T_i . Such i exist because otherwise all matrices in the image (in particular, A and B) have the same ratio of eigenvalues. But $\frac{(\text{tr } f)^2}{\det f}$ is the ratio of quadratic polynomials, and K is quadratically closed.

If there is a point T_i such that $\text{tr } f = \det f = 0$, then f evaluated at this T_i is nilpotent. Since $\text{tr } f$ is a linear function, the equation $\text{tr } f = 0$ has only one root, which is a rational function on the other parameters. Thus f evaluated at this T_i is 0, by Amitsur's theorem. We conclude that the ratio of eigenvalues does not depend on T_i , contrary to our assumption on i . Hence, we can solve $\frac{(\text{tr } f)^2}{\det f} = c$ for any $c \in K$. ■

Lemma 2.11. *If there exist $\lambda_1 \neq \pm \lambda_2$ with a collection of matrices (A_1, A_2, \dots, A_m) such that $p(A_1, A_2, \dots, A_m)$ has eigenvalues λ_1 and λ_2 , then all diagonalizable matrices lie in $\text{Im } p$.*

Proof. Applying Lemma 1.27 to the hypothesis, there is a matrix

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \in \text{Im } p, \quad \lambda_1 \neq \pm \lambda_2$$

which is an evaluation of p on matrix units e_{ij} . Consider the following mapping χ acting on the indices of the matrix units: $\chi(e_{ij}) = e_{3-i, 3-j}$. Now take the polynomial

$$f(T_1, T_2, \dots, T_m) = p(\tau_1 x_1 + t_1 \chi(x_1), \dots, \tau_m x_m + t_m \chi(x_m)),$$

where $T_i = (t_i, \tau_i) \in K^2$, which is linear with respect to each T_i . Let us open the brackets. We obtain 2^m terms and for each of them the degrees of all vertices stay even. (The edge 12 becomes 21 which does not change degrees, and the edge 11 becomes 22, which decreases the degree of the vertex 1 by two and increases the degree of the vertex 2 by two.) Thus all terms remain diagonal. Consider generic pairs $T_1, \dots, T_m \in K^2$. For each i consider the polynomial $\tilde{f}_i(T_i^*) = f(T_1, \dots, T_{i-1}, T_i + T_i^*, T_{i+1}, \dots, T_m)$. For at least one i the ratio of eigenvalues of \tilde{f}_i must be different from ± 1 . (Otherwise the ratio of eigenvalues of \tilde{f}_i equal ± 1 all i , implying $\lambda_1 = \pm \lambda_2$ }, a contradiction.)

Fix i such that the ratio of eigenvalues of \tilde{f}_i is not ± 1 . By linearity, $\text{Im}(\tilde{f}_i)$ takes on values with all possible ratios of eigenvalues; hence, the cone under $\text{Im}(\tilde{f}_i)$ is the set of all diagonal matrices. Therefore by Lemma 1.46 all diagonalizable matrices lie in the image of p . ■

Images of multilinear polynomials.

Lemma 2.12. *If p is a multilinear polynomial evaluated on the matrix ring $M_2(K)$ over a quadratically closed field K , then $\text{Im } p$ is either $\{0\}$, K , sl_2 , $M_2(K)$, or $M_2(K) \setminus \tilde{K}$.*

Proof. If $\text{Im } p$ does not contain a non-scalar matrix, then p is either PI or central, and we are done. Hence, without loss of generality we can assume that $\text{Im } p$ contains a non-scalar matrix. By Remark 2.7 the linear span of $\text{Im } p$ is sl_2 or $M_2(K)$. We treat the characteristic 2 and characteristic $\neq 2$ cases separately.

Case I: Char $K = 2$. Consider the set

$$\Theta = \{p(e_1, \dots, e_m) \text{ where the } e_j \text{ are matrix units}\}.$$

If the linear span of the image is not sl_2 , then Θ contains at least one non-scalar diagonal matrix $\text{Diag}\{\lambda_1, \lambda_2\}$, so $\lambda_1 \neq -\lambda_2$ (since $+1 = -1$). Hence by Lemma 2.11, all diagonalizable matrices belong to $\text{Im } p$. Thus, $\text{Im } p$ contains $M_2(K) \setminus \tilde{K}$.

If the linear span of the image of p is sl_2 , then by Lemma 1.27 the identity matrix (and thus all scalar matrices) and e_{12} (and thus all nilpotent matrices) belong to the image. On the other hand, in characteristic 2, any matrix sl_2 is conjugate to a matrix of the form $\lambda_1 I + \lambda_2 e_{1,2}$, and we consider the invariant $\frac{\lambda_2}{\lambda_1}$. Take x_1, \dots, x_m to be generic matrices. If $p(x_1, \dots, x_m)$ were nilpotent then $\text{Im } p$ would consist only of nilpotent matrices, which is impossible. By Example 2.2(v), $p(x_1, \dots, x_m)$ is not scalar and not nilpotent, and thus is a matrix from \tilde{K} . Hence, $\tilde{K} \subset \text{Im } p$, by Remark 2.7. Thus, all trace zero matrices belong to $\text{Im } p$.

Case II: Char $K \neq 2$. Again assume that the image is not $\{0\}$ or the set of scalar matrices. Then $e_{12} \in \text{Im } p$ by Lemma 1.27. Thus all nilpotent matrices lie in $\text{Im } p$. If the image consists only of matrices of trace zero, then there is at least one matrix in the image with a nonzero diagonal entry. By Lemma 1.27 there is a set of matrix units that maps to a nonzero diagonal matrix which, by assumption, is of trace zero and thus is $\begin{pmatrix} c & 0 \\ 0 & -c \end{pmatrix}$. By Lemma 1.46 and Example 2.2, $\text{Im } p$ contains all trace zero 2×2 matrices.

Assume that the image contains a matrix with nonzero trace. Then by Remark 2.7 the linear span of the image is $M_2(K)$, and together with Lemma 1.27 we have at least two diagonal linearly independent matrices in the image. Either these matrices have ratios of eigenvalues $(\lambda_1 : \lambda_2)$ and $(\lambda_2 : \lambda_1)$ for $\lambda_1 \neq \pm \lambda_2$ or these matrices have non-equivalent ratios. In the first case we can use Lemma 2.11 which says that all diagonalizable matrices lie in the image. If at least one of these matrices have ratio not equal to ± 1 , then in the second case we also use Lemma 2.11 and obtain that all diagonalizable matrices lie in the image. If these matrices are such that the ratios of their eigenvalues are respectively 1 and -1 , then we use Lemma 2.10 and obtain that all diagonalizable matrices with distinct eigenvalues lie in the image. By assumption, in this case, scalar matrices also belong to the image. Therefore we obtain that for any ratio $(\lambda_1 : \lambda_2)$ there is a matrix $A \in \text{Im } p$ having such a ratio of eigenvalues. Using Lemmas 1.46 and 2.7, we obtain that the image of p can be either $\{0\}$, K , sl_2 , $M_2(K)$, or $M_2(K) \setminus \tilde{K}$. ■

Lemma 2.13. *If p is a multilinear polynomial evaluated on the matrix ring $M_2(K)$, where K is a quadratically closed field of characteristic 2, then $\text{Im } p$ is either $\{0\}$, K , sl_2 , or $M_2(K)$.*

Proof. In view of Lemma 2.12, it suffices to assume that the image of p is $M_2(K) \setminus \tilde{K}$. Let $x_1, \dots, x_m, y_1, \dots, y_m$ be generic matrices. Consider the polynomials

$$b_i = p(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_m).$$

Let $p_i(x_1, \dots, x_m, y_i) = p \text{tr}(b_i) + \text{tr}(p)b_i$. Hence p_i can be written as

$$p_i = p(x_1, \dots, x_{i-1}, x_i \text{tr}(b_i) + y_i \text{tr}(p), x_{i+1}, \dots, x_m).$$

Therefore $\text{Im } p_i \subseteq \text{Im } p$. Also if $a \in \text{Im } p_i$, then

$$\text{tr}(a) = \text{tr}(p \text{tr}(b_i) + \text{tr}(p)b_i) = 2 \text{tr}(p) \text{tr}(b_i) = 0.$$

Thus, $\text{Im } p_i$ consists only of trace-zero matrices which belong to the image of p . Excluding \tilde{K} , the only trace zero matrices are nilpotent or scalar. Thus, for each i , $p_i(x_1, \dots, x_m, y_i)$ is either scalar or nilpotent. However, the p_i are the elements of the algebra of generic matrices with traces, which is a domain. Thus, $p_i(x_1, \dots, x_m, y_i)$ cannot be nilpotent. Hence for all $i = 1, \dots, m$, $p_i(x_1, \dots, x_m, y_i)$ is scalar. In this case, changing variables leaves the plane $\langle p, I \rangle$ invariant. Therefore, $\dim(\text{Im } p) = 2$, a contradiction. \blacksquare

Lemma 2.14. *If p is a multilinear polynomial evaluated on the matrix ring $M_2(K)$ (where K is a quadratically closed field of characteristic not 2), then $\text{Im } p$ is either $\{0\}$, K , sl_2 , or $M_2(K)$.*

The proof of Lemma 2.14 is given in Section 2.2.

Finally, Theorem 2.4 follows from Lemmas 2.13 and 2.14.

2.1.2 Images of arbitrary non-homogeneous polynomials

We consider briefly the general non-homogeneous case. One can write any polynomial $p(x_1, \dots, x_m)$ as $p = h_k + \dots + h_n$, where the h_i are semi-homogeneous polynomials of weighted degree i .

Proposition 2.15. *Notation as above, assume that there are weights (w_1, \dots, w_m) that $\text{Im } h_n$ is dense in $M_2(K)$. Then $\text{Im } p$ is dense in $M_2(K)$.*

Proof. Consider

$$p(\lambda_1^w x_1, \dots, \lambda_m^w x_m) = \sum_{i=k}^n h_i \lambda^i.$$

One can write $\tilde{P} = \lambda^{-n} p(\lambda_1^w x_1, \dots, \lambda_m^w x_m)$ as a polynomial in x_1, \dots, x_m and $\varepsilon = \frac{1}{\lambda}$. The matrix polynomial is the set of four polynomials $p_{1,1}, p_{1,2}, p_{2,1}, p_{2,2}$, which we claim are independent. If there is some polynomial h in four variables such that $h(p_{1,1}, p_{1,2}, p_{2,1}, p_{2,2}) = 0$ then h should vanish on four polynomials of \tilde{P} for each ε , in particular for $\varepsilon = 0$, a contradiction. \blacksquare

Remark 2.16. The case remains open where $p(x_1, \dots, x_m)$ is a polynomial for which there are no weights (w_1, \dots, w_m) such that one can write $p = h_k + \dots + h_n$, where h_i is semi-homogeneous of weighted degree i and h_n has dense image in M_2 .

Example 2.17. For $\text{Char } K \neq 2$ we give an example of such a polynomial whose middle term has dense image in $M_2(K)$. Take the polynomial

$$f(x, y) = [x, y] + [x, y]^2.$$

It is not hard to check that $\text{Im } f$ is the set of all matrices with eigenvalues $c^2 + c$ and $c^2 - c$. Consider $p(\alpha_1, \alpha_2, \beta_1, \beta_2) = f(\alpha_1 + \beta_1^2, \alpha_2 + \beta_2^2)$. The polynomials f and p have the same images. Now let us open the brackets. The term of degree 4 is $h_4 = [\alpha_1, \alpha_2]^2 + [\beta_1^2, \beta_2^2]$. The image of h_4 is all of $M_2(K)$, because $[\alpha_1, \alpha_2]^2$ can be any scalar matrix and $[\beta_1^2, \beta_2^2]$ can be any trace zero matrix. However the image of p is the set of all matrices with eigenvalues $c^2 + c$ and $c^2 - c$.

2.1.3 The cases of an arbitrary real field

Let p be a multilinear polynomial in several non-commuting variables with coefficients in an arbitrary field K . Kaplansky conjectured that for any n , the image of p evaluated on the set $M_n(K)$ of n by n matrices is either zero, or the set of scalar matrices, or the set $\text{sl}_n(K)$ of matrices of trace 0, or all of $M_n(K)$. This conjecture was proved for $n = 2$ when K is closed under quadratic extensions. In this section the conjecture is verified for $K = \mathbb{R}$ and $n = 2$, also for semi-homogeneous polynomials p , with a partial solution for an arbitrary field K .

In Section 2.1.1 the field K was required to be quadratically closed. Even for the field \mathbb{R} of real numbers L'vov–Kaplansky's question remained open, leading people to ask what happens if the field is not quadratically closed? This subsection provides a positive partial answer.

2.1.4 Images of multilinear polynomials evaluated on $M_2(K)$

Assume that p is a multilinear polynomial evaluated on 2×2 matrices over any field K . Assume also that p is neither PI nor central. Then, by Lemma 1.27 there exist matrix units a_1, \dots, a_m such that $p(a_1, \dots, a_m) = e_{12}$. Let us consider the mapping χ defined on matrix units that switches the indices 1 and 2, i.e., $e_{11} \leftrightarrow e_{22}$ and $e_{12} \leftrightarrow e_{21}$. Now let us consider the mapping f defined on m pairs $T_i = (t_i, \tau_i)$:

$$f(T_1, \dots, T_m) = p(t_1 a_1 + \tau_1 \chi(a_1), t_2 a_2 + \tau_2 \chi(a_2), \dots, t_m a_m + \tau_m \chi(a_m)).$$

Now let us open the brackets. We show in the proof of Theorem 1.12 that any matrix of the image of f can be written as $c_1 e_{12} + \dots + c_{n-1} e_{n-1, n} + c_n e_{n, n-1}$. In our case $n = 2$ and the image of f contains only matrices of the type $c_1 e_{12} + c_2 e_{21}$. Note that the matrices e_{12} and e_{21} both belong to the image of f since $p(a_1, \dots, a_m) = e_{12}$ and $p(\chi(a_1), \dots, \chi(a_m)) = e_{21}$. According to Lemma 1.34 the image of f is at least 2-dimensional, and lies in the 2-dimensional plane $\langle e_{12}, e_{21} \rangle$. Therefore this plane is exactly the image of f . Now we are ready to prove the following:

Lemma 2.18. *If p is a multilinear polynomial evaluated on the matrix ring $M_2(K)$ (for an arbitrary field K), then $\text{Im } p$ is either $\{0\}$ or K , or $\text{sl}_2 \setminus K \subseteq \text{Im } p$.*

Proof. Let A be any trace zero, non-scalar matrix. Take any vector v_1 that is not an eigenvector of A . Consider the vector $v_2 = Av_1$. Note that

$$Av_2 = A^2 v_1 = -\det(A)v_1,$$

and therefore the matrix A with respect to the base $\{v_1, v_2\}$ has the form $c_1 e_{12} + c_2 e_{21}$, for some c_i . Hence A is similar to $c_1 e_{12} + c_2 e_{21} \in \text{Im } p$, implying $A \in \text{Im } p$. ■

Remark 2.19. Note that for $\text{Char}(K) \neq 2$ (in particular for $K = \mathbb{R}$),

$$(\text{sl}_2 \setminus K) \cup \{0\} = \text{sl}_2 \subseteq \text{Im } p.$$

The real case. Now we assume that $K = \mathbb{R}$. We already know that either p is PI, or central, or $\text{sl}_2 \subseteq \text{Im } p$. Assume that $\text{sl}_2 \subsetneq \text{Im } p$. We will use the following lemma:

Lemma 2.20. *Let p be any multilinear polynomial satisfying $\text{sl}_2 \subsetneq \text{Im } p$. For any $q \in \mathbb{R}$ there exist generic matrices $x_1, \dots, x_m, y_1, \dots, y_m$ such that for $X = p(x_1, \dots, x_m)$ and $Y = p(y_1, \dots, y_m)$ we have the following:*

$$\frac{\det X}{\text{tr}^2 X} \leq q \leq \frac{\det Y}{\text{tr}^2 Y},$$

where $\text{tr}^2 M$ denotes the square of the trace of M .

Proof. We know that $\mathfrak{sl}_2 \subseteq \text{Im } p$, in particular for the matrices $\Omega = e_{11} - e_{22}$ and $\Upsilon = e_{12} - e_{21}$ there exist matrices $a_1, \dots, a_m, b_1, \dots, b_m$ such that $p(a_1, \dots, a_m) = \Omega$ and $p(b_1, \dots, b_m) = \Upsilon$. Note $\frac{\det M}{\text{tr}^2 M} \leq q$ if M is close to Ω and $\frac{\det M}{\text{tr}^2 M} > q$ if M is close to Υ . Now we consider a very small $\delta > 0$ such that for any matrices $x_i \in N_\delta(a_i)$ and $y_i \in N_\delta(b_i)$

$$\frac{\det X}{\text{tr}^2 X} \leq q \leq \frac{\det Y}{\text{tr}^2 Y},$$

where $X = p(x_1, \dots, x_m)$ and $Y = p(y_1, \dots, y_m)$. Here by $N_\delta(x)$ we denote a δ -neighborhood of x , under the max norm $\|A\| = \max_{i,j} |a_{ij}|$. According to Lemma 1.40 one can choose generic matrices with such property. \blacksquare

Now we are ready to prove that the image of $g(x_1, \dots, x_m) = \frac{\det p}{\text{tr}^2 p}$ is all of \mathbb{R} :

Lemma 2.21. *Let p be any multilinear polynomial satisfying $\mathfrak{sl}_2 \subsetneq \text{Im } p$. Then for any $q \in \mathbb{R}$ there exists a set of matrices a_1, \dots, a_m such that*

$$\frac{\det p(a_1, \dots, a_m)}{\text{tr}^2 p(a_1, \dots, a_m)} = q. \quad (2.1)$$

Proof. Let q be any real number. According to Lemma 2.20 there exist generic matrices $x_1, \dots, x_m, y_1, \dots, y_m$ such that for $X = p(x_1, \dots, x_m)$ and $Y = p(y_1, \dots, y_m)$ we have the following:

$$\frac{\det X}{\text{tr}^2 X} \leq q \leq \frac{\det Y}{\text{tr}^2 Y}.$$

Consider the following matrices: $A_0 = p(\tilde{x}_1, x_2, \dots, x_m)$, where \tilde{x}_1 is either x_1 or $-x_1$, such that $\text{tr } A_0 > 0$. $A_1 = p(\tilde{y}_1, x_2, \dots, x_m)$, where \tilde{y}_1 is either y_1 or $-y_1$ such that $\text{tr } A_1 > 0$. Assume that $A_i, \tilde{x}_1, \tilde{y}_1, \dots, \tilde{y}_i$ are defined. Let

$$A_{i+1} = p(\tilde{y}_1, \dots, \tilde{y}_i, \tilde{y}_{i+1}, x_{i+2}, \dots, x_m),$$

where $\tilde{y}_{i+1} = \pm y_{i+1}$ is such that $\text{tr } A_{i+1} > 0$. In such a way we defined matrices A_i for $0 \leq i \leq m$. Note that for any 2×2 matrix M ,

$$\frac{\det M}{\text{tr}^2 M} = \frac{\det(-M)}{\text{tr}^2(-M)}.$$

Note that $A_0 = \pm p(x_1, \dots, x_m)$ and $A_m = \pm p(y_1, \dots, y_m)$; hence

$$\frac{\det A_0}{\text{tr}^2 A_0} \leq q \leq \frac{\det A_m}{\text{tr}^2 A_m}.$$

Therefore there exists i such that

$$\frac{\det A_i}{\text{tr}^2 A_i} \leq q \leq \frac{\det A_{i+1}}{\text{tr}^2 A_{i+1}}.$$

Since $A_i = p(\tilde{y}_1, \dots, \tilde{y}_i, x_{i+1}, x_{i+2}, \dots, x_m)$ and $A_{i+1} = p(\tilde{y}_1, \dots, \tilde{y}_{i+1}, x_{i+2}, \dots, x_m)$, we can consider the matrix function

$$M(t) = (1-t)A_i + tA_{i+1} = p(\tilde{y}_1, \dots, \tilde{y}_i, (1-t)x_{i+1} + t\tilde{y}_{i+1}, x_{i+2}, \dots, x_m).$$

Then $\text{Im } M \subseteq \text{Im } p$, $M(0) = A_i$, $M(1) = A_{i+1}$ both $M(0)$ and $M(1)$ have positive trace, and M is an affine function. Therefore for any $t \in [0, 1]$ $M(t)$ has positive trace. Therefore the function $\psi(t) = \frac{\det M(t)}{\text{tr}^2 M(t)}$ is well defined on $[0, 1]$ and continuous. Also we have $\psi(0) \leq q \leq \psi(1)$. Thus there exists $\tau \in [0, 1]$ such that $\psi(\tau) = q$ and thus $M(\tau) \in \text{Im } p$ satisfies equation (2.1). \blacksquare

Lemma 2.22. *Let p be a multilinear polynomial satisfying $\text{sl}_2 \not\subseteq \text{Im } p$. Then any matrix with distinct eigenvalues (i.e., matrix of nonzero discriminant) belongs to $\text{Im } p$.*

Proof. Let A be any matrix with nonzero discriminant. Let us show that $A \in \text{Im } p$. Let $q = \frac{\det A}{\text{tr}^2 A}$. According to Lemma 2.21 there exists a set of matrices a_1, \dots, a_m such that $\frac{\det \tilde{A}}{\text{tr}^2 \tilde{A}} = q$, where $\tilde{A} = p(a_1, \dots, a_m)$. Take $c \in \mathbb{R}$ such that $\text{tr}(c\tilde{A}) = \text{tr } A$. Note $c\tilde{A} = p(ca_1, a_2, \dots, a_m)$ belongs to $\text{Im } p$. Thus

$$\frac{\det(c\tilde{A})}{\text{tr}^2(c\tilde{A})} = q = \frac{\det A}{\text{tr}^2 A},$$

and $\text{tr } A = \text{tr}(c\tilde{A})$. Hence, $\det(c\tilde{A}) = \det(A)$. Therefore the matrices $c\tilde{A}$ and A are similar since they are not from the discriminant surface. Therefore $A \in \text{Im } p$. ■

Lemma 2.23. *Let p be a multilinear polynomial satisfying $\text{sl}_2 \not\subseteq \text{Im } p$. Then any non-scalar matrix with zero discriminant belongs to $\text{Im } p$.*

Proof. Let A be any non-scalar matrix with zero discriminant. Let us show that $A \in \text{Im } p$. The eigenvalues of A are equal, and therefore they must be real. Thus A is similar to the matrix $\tilde{A} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. If A is nilpotent then $\lambda = 0$ and $\tilde{A} = e_{12}$, and it belongs to $\text{Im } p$ by Lemma 1.27. If A is not nilpotent then we need to show that at least one non-nilpotent matrix of such type belongs to $\text{Im } p$, and all other are similar to it. We know that the matrices $e_{11} - e_{22} = p(a_1, \dots, a_m)$ and $e_{12} - e_{21} = p(b_1, \dots, b_m)$ for some a_i and b_i . Note that $e_{11} - e_{22}$ has positive discriminant and $e_{12} - e_{21}$ has negative discriminant. Take generic matrices $x_1, x_2, \dots, x_m, y_1, \dots, y_m$ such that $x_i \in N_\delta(a_i)$ and $y_i \in N_\delta(b_i)$ where $\delta > 0$ is so small that $p(x_1, \dots, x_m)$ has positive discriminant and $p(y_1, \dots, y_m)$ has negative discriminant. Consider the following matrices:

$$A_0 = p(x_1, x_2, \dots, x_m), \quad A_i = p(y_1, \dots, y_i, x_{i+1}, \dots, x_m), \quad 1 \leq i \leq m.$$

We know that $\text{Discr } A_0 > 0$ and $\text{Discr } A_m < 0$, and therefore there exists i such that $\text{Discr } A_i > 0$ and $\text{Discr } A_{i+1} < 0$. We can consider the continuous matrix function

$$M(t) = (1-t)A_i + tA_{i+1} = p(y_1, \dots, y_i, (1-t)x_{i+1} + ty_{i+1}, x_{i+2}, \dots, x_m).$$

We know that $M(0)$ has positive discriminant and $M(1)$ has negative discriminant. Therefore for some τ , $M(\tau)$ has discriminant zero. Assume there exists t such that $M(t)$ is nilpotent. In this case either t is unique or there exists $t' \neq t$ such that $M(t')$ is also nilpotent. If t is unique then it equals to some rational function with respect to other variables (entries of matrices x_i and y_i). In this case t can be considered as a function on matrices $\widetilde{x_i}$ and $\widetilde{y_i}$ and as soon as it is invariant, according to the Proposition 1.30 t is an element of $\widetilde{\text{UD}}$ and thus $M(t)$ is the element of $\widetilde{\text{UD}}$. Therefore $M(t)$ cannot be nilpotent since UD is a domain according to Remark 1.31. If there exists $t' \neq t$ such that $M(t')$ is also nilpotent then for any $\tilde{t} \in \mathbb{R}$ $M(\tilde{t})$ is the combination of two nilpotent (and thus trace vanishing) matrices $M(t)$ and $M(t')$. Hence $M(0)$ is trace vanishing and thus $\text{Im } p \subseteq \text{sl}_2$, a contradiction.

Recall that we proved $M(\tau)$ has discriminant zero that for some τ . Note that $M(\tau)$ cannot be nilpotent. Assume that the matrix $M(\tau)$ is scalar. Hence $(1-\tau)A_i + \tau A_{i+1} = \lambda I$ where $\lambda \in \mathbb{R}$ and I is the identity matrix. Thus, $A_{i+1} = \frac{1-\tau}{\tau} A_i + cI$. Note that for any matrix M and any $c \in \mathbb{R}$ we have $\text{Discr}(M) = \text{Discr}(M + cI)$. Therefore the discriminant of A_{i+1} can be written as

$$\text{Discr}(A_{i+1}) = \text{Discr}\left(\frac{1-\tau}{\tau} A_i\right) = \left(\frac{1-\tau}{\tau}\right)^2 \text{Discr}(A_i),$$

a contradiction, since $\text{Discr } A_i > 0$ and $\text{Discr}(A_{i+1}) < 0$. Therefore the matrix $M(\tau)$ is similar to A . ■

Lemma 2.24. *Let p be a multilinear polynomial satisfying $\mathfrak{sl}_2 \subsetneq \text{Im } p$. Then every scalar matrix belongs to $\text{Im } p$.*

Proof. Note that it is enough to show that at least one scalar matrix belong to the image of p . According to Lemmas 1.27 and Remark 2.7 there are matrix units a_1, \dots, a_m such that $p(a_1, \dots, a_m)$ is diagonal with nonzero trace. Assume that it is not scalar, i.e., $p(a_1, \dots, a_m) = \lambda_1 e_{11} + \lambda_2 e_{22}$. We define again the mapping χ and $f(T_1, \dots, T_m)$ as in the beginning of Section 2.1.4 and return to the proof of Lemma 2.11 where we proved that $\text{Im } f$ consists only of diagonal matrices or only of matrices with zeros on the diagonal. In our case the image of f consists only of diagonal matrices, which is a 2-dimensional variety. We know that both $p(a_1, \dots, a_m) = \lambda_1 e_{11} + \lambda_2 e_{22}$ and $p(\chi(a_1), \dots, \chi(a_m)) = \lambda_1 e_{22} + \lambda_2 e_{11}$ belong to the image of f , and therefore every diagonal matrix belong to the image of f , in particular every scalar matrix. ■

Now we are ready to prove a major result.

Proof of Theorem 1.15. The second part follows from Lemmas 2.18, 2.22, 2.23 and 2.24. In the first part we need to prove that if p is neither PI nor central then $\mathfrak{sl}_2(K) \subseteq \text{Im } p$. According to Lemma 2.18, $\mathfrak{sl}_2(K) \setminus K \subseteq \text{Im } p$, and therefore according to Remark 2.19 we need consider only the case $\text{Char}(K) = 2$. Then we need to prove that the scalar matrices belong to the image of p . By Lemma 1.27 and Remark 2.7 there are matrix units a_1, \dots, a_m such that $p(a_1, \dots, a_m)$ is diagonal. Assume that it is not scalar. Then we consider the mappings χ and f as described in the beginning of Section 2.1.4. According to Lemma 1.34 the image of f will be the set of all diagonal matrices, and in particular the scalar matrices belong to it. ■

Remark 2.25. Assume that p is a multilinear polynomial evaluated on 2×2 matrices over an arbitrary infinite field K . Then, according to Theorem 1.15, $\text{Im } p$ is $\{0\}$, or K , or $\mathfrak{sl}_2(K)$ or $\mathfrak{sl}_2(K) \subsetneq \text{Im } p$. In the last case it is clear that $\text{Im } p$ must be Zariski dense in $M_2(K)$, because otherwise $\dim(\text{Im } p) = 3$ and $\text{Im } p$ is reducible, a contradiction.

Remark 2.26. Note that the proof of Theorem 1.15 does not work when $n > 2$ since for this case we will need to take more than one function (two functions for $n = 3$ and more for $n > 3$). In our proof we used that we have only one function: we proved that it takes values close to $\pm\infty$ and after that used continuity. This does not work for $n \geq 3$. However one can use this idea for the question of possible images of trace vanishing multilinear polynomials evaluated on 3×3 matrices. In this case one function will be enough, and one can take $g = \frac{\omega_2^2}{\omega_3^2}$. (One can find the definitions of ω_i in the proof of Theorem 1.18.) Moreover by Lemma 1.27 there are matrix units a_i such that $p(a_1, \dots, a_m)$ is a diagonal, trace vanishing, nonzero real matrix, which cannot be 3-scalar since it will have three real eigenvalues. Therefore p cannot be 3-central polynomial. However the question of possible images of p remains open.

Images of semi-homogeneous polynomials evaluated on 2×2 matrices with real entries. Here we provide a classification of the possible images of semi-homogeneous polynomials evaluated on 2×2 matrices with real entries. Let us start with the definitions.

Definition 2.27. A *semi-cone* of $M_n(\mathbb{R})$ is a subset closed under multiplication by positive constants. An *invariant semi-cone* is a semi-cone invariant under conjugation. An invariant semi-cone is *irreducible* if it does not contain any nonempty invariant semi-cone.

Remark 2.28. Let p be any semi-homogeneous polynomial of weighted degree $d \neq 0$ with weights (w_1, \dots, w_m) . Thus if $A = p(x_1, \dots, x_m)$ then for any $c \in \mathbb{R}$ we have $p(c^{w_1}x_1, \dots, c^{w_m}x_m) = c^d A$. Hence $\text{Im } p$ is a semi-cone, for any d .

Proof of Theorem 1.17. Consider the function $g(x_1, \dots, x_m) = \frac{\det p}{\operatorname{tr}^2 p}$. If this function is not constant, then $\operatorname{Im} p$ is Zariski dense. Assume that it is constant; i.e., $\frac{\det p}{\operatorname{tr}^2 p} = c$. Then the ratio $\frac{\lambda_1}{\lambda_2} = \hat{c}$ of eigenvalues is also a constant. If $\hat{c} \neq -1$ then we can write λ_1 explicitly as

$$\lambda_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2} \operatorname{tr} p = \frac{1}{1 + \frac{\lambda_2}{\lambda_1}} \operatorname{tr} p = \frac{1}{1 + \frac{1}{\hat{c}}} \operatorname{tr} p,$$

Therefore λ_1 is an element of $\widetilde{\operatorname{UD}}$, and $\lambda_2 = \operatorname{tr} p - \lambda_1$ also. According to the Hamilton–Cayley equation, $(p - \lambda_1)(p - \lambda_2) = 0$ and therefore, since, by Remark 1.31, $\widetilde{\operatorname{UD}}$ is a domain, one of the terms $p - \lambda_i$ is a PI. Therefore p is central or PI. Therefore we see that any semi-homogeneous polynomial is either PI, or central, or trace vanishing (if the ratio of eigenvalues is -1 then the trace is identically zero), or $\operatorname{Im} p$ is Zariski dense. If p is PI then $\operatorname{Im} p = \{0\}$. If p is central then, by Remark 2.28, $\operatorname{Im} p$ is a semi-cone, therefore $\operatorname{Im} p$ is either $\mathbb{R}_{\geq 0}$, or $\mathbb{R}_{\leq 0}$, or \mathbb{R} . If p is trace vanishing, then any trace zero matrix $A \in \operatorname{sl}_2(\mathbb{R})$ is similar to $-A$. Therefore $\operatorname{Im} p = -\operatorname{Im} p$ is symmetric. Together with Remark 2.28 we have that $\operatorname{Im} p$ must be a cone. The determinant cannot be identically zero since otherwise the polynomial is nilpotent, contrary to Remark 1.31. Hence there exists some value with nonzero determinant. All the trace zero matrices of positive determinant are pairwise similar, and all the trace zero matrices of negative determinant are pairwise similar. Therefore in this case all possible images of p are $\operatorname{sl}_{2, \geq 0}(\mathbb{R})$, $\operatorname{sl}_{2, \leq 0}(\mathbb{R})$ and $\operatorname{sl}_2(\mathbb{R})$. ■

Example 2.29. $\operatorname{Im} p$ can be the set of non-negative scalars. Take any central polynomial, say $p(x, y) = [x, y]^2$ and consider $p^2 = [x, y]^4$. If one takes $-p^2 = -[x, y]^4$, then its image is the set $\mathbb{R}_{\leq 0}$.

The question remains open of whether or not there exists an example of a trace zero polynomial with non-negative (or non-positive) discriminant.

There are many polynomials with Zariski dense image which are not dense with respect to the usual Euclidean topology. For example the image of the polynomial $p(x) = x^2$ is the set of matrices with two positive eigenvalues, or two complex conjugate eigenvalues; in particular any matrix x^2 has non-negative determinant. The image of the polynomial $p(x, y) = [x, y]^4 + [x^4, y^4]$ is the set of matrices with non-negative trace. The question of classifying possible semi-homogeneous Zariski dense images remains open.

2.1.5 Multilinear polynomials evaluated on \mathbb{H}

The L’vov–Kaplansky conjecture can fail for a non-simple finite dimensional algebra. In particular, it fails for the Grassmann algebra over a linear space of finite dimension more than or equal to 4, and a field of characteristic $\neq 2$, which is a finite dimensional (but not simple) associative algebra. In this case one can consider the multilinear polynomial $p(x, y) = x \wedge y - y \wedge x$. Then $e_1 \wedge e_2 = p(1/2e_1, e_2)$ and $e_3 \wedge e_4$ both belong to the image of p , but their sum does not. Thus the image is not a vector space.

Recently, it has been conjectured that the evaluation of any multilinear polynomial on a simple algebra is a vector space. However, according to [80] this conjecture fails even for some (infinite dimensional) division algebras and the polynomial $p(x, y) = xy - yx$. Cohn [24] constructed a division ring D in which every element is a commutator, i.e., $p(D) = D$.

It is interesting to investigate the Kaplansky conjecture for finite dimensional simple algebras. In this section we deal with quaternions. Note that the algebra of split quaternions, defined also by 4-dimensional vector space $\langle 1, i, j, k \rangle_{\mathbb{R}}$ with multiplication defined by

$$ij = -ji = k, \quad jk = -kj = -i, \quad ki = -ik = j, \quad i^2 = -1, \quad j^2 = k^2 = 1,$$

is isomorphic to $M_2(\mathbb{R})$. Let us start with proving the following straightforward but important lemmas:

Lemma 2.30. *Let p be a multilinear polynomial. If a_i are basic quaternions $(1, i, j, k)$, then $p(a_1, \dots, a_m)$ is $c \cdot q$ for some basic quaternion q and some scalar $c \in \mathbb{R}$ (which can equal 0).*

Proof. Note that for any two basic quaternions q_1 and q_2 , $q_1 q_2 = \pm q_2 q_1$. Therefore, taking products of m basic quaternions we obtain the same result (up to \pm). Thus the sum of these results multiplied by scalars must be a basic quaternion multiplied by some scalar coefficient. ■

Lemma 2.31. *For any multilinear polynomial p , $\text{Im } p$ is a self-similar cone, i.e., for any invertible $h \in \mathbb{H}$, any scalar $c \in \mathbb{R}$ and any element $\alpha \in \text{Im } p$, $ch\alpha h^{-1} \in \text{Im } p$.*

Proof. If $p(x_1, \dots, x_m) = \sum_{\sigma \in S_m} c_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)}$, then

$$p(hx_1 h^{-1}, hx_2 h^{-1}, \dots, hx_m h^{-1}) = \sum_{\sigma \in S_m} c_\sigma h x_{\sigma(1)} h^{-1} \cdots h x_{\sigma(m)} h^{-1} = h p(x_1, \dots, x_m) h^{-1},$$

and thus $p(chx_1 h^{-1}, hx_2 h^{-1}, \dots, hx_m h^{-1}) = ch\alpha h^{-1} \in \text{Im } p$. ■

Lemma 2.32. *The set of pure quaternions V is an irreducible self-similar cone, i.e., equals all of its conjugates.*

Proof. It is enough to show that any self-similar cone including the element i contains V . Take $h(y, z) = 1 + yj + zk$, thus $h^{-1} = \frac{1-yj-zk}{1+y^2+z^2}$. Thus a minimal self-similar cone C containing i contains all elements $c \cdot hih^{-1}$, in particular it contains all elements $(1 + yj + zk)i(1 - yj - zk) = (1 - y^2 - z^2)i + 2zj - 2yk$. Consider an arbitrary pure quaternion $ai + bj + ck$. If $b = c = 0$ then this vector belongs to C because it is a multiple of i . Assume that at least one of b and c is nonzero. Then $b^2 + c^2 > 0$, hence one can take $l = \frac{a + \sqrt{a^2 + b^2 + c^2}}{2}$, $y = -\frac{c}{a + \sqrt{a^2 + b^2 + c^2}}$, $z = \frac{b}{a + \sqrt{a^2 + b^2 + c^2}}$, these numbers are well defined. Thus the element $(1 + yj + zk)i(1 - yj - zk) = (1 - y^2 - z^2)i + 2zj - 2yk = ai + bj + ck$ belongs to C . ■

Now we are ready to prove the main theorem:

Proof of Theorem 1.25. Let us substitute basic quaternions in the polynomial p . Lemma 2.30 implies that we obtain multiples of basic quaternions. Consider the four possible cases: all these results vanish, or among these results there are scalars only, pure quaternions only, and both pure quaternions and scalars. The first two cases quickly lead to answers about the image of the polynomial p : in the first case p is PI, its image is $\{0\}$, and in the second case it is central polynomial, and its image is \mathbb{R} . In the third case the image is V , by Lemma 2.32.

Therefore the most interesting case is the fourth one. We assume that there are basic quaternions x_1, \dots, x_m and y_1, \dots, y_m such that $p(x_1, \dots, x_m) = k \in \mathbb{R} \setminus \{0\}$ and $p(y_1, \dots, y_m) = v \in V \setminus \{0\}$. We will show that in this case the image of p is the set of all the quaternions. For that let us consider the following $m + 1$ evaluations depending on y_1, \dots, y_m : $A_1 = p(x_1, x_2, \dots, x_m)$; $A_2 = p(y_1, x_2, \dots, x_m)$; $A_3 = p(y_1, y_2, x_3, \dots, x_m)$; \dots ; $A_{m+1} = p(y_1, \dots, z_m)$. Note that A_1 is a constant taking only one possible value (which is a nonzero scalar), for any i $\text{Im } A_i \subseteq \text{Im } A_{i+1}$ and $\text{Im } A_{m+1} = \text{Im } p$ includes nonscalar values. Therefore, there exists i such that $\text{Im } A_i \subseteq \mathbb{R}$ and $\text{Im } A_{i+1} \not\subseteq \mathbb{R}$. Thus there exist some collection of quaternions $r_1, r_2, \dots, r_m, r_i^*$ such that $p(r_1, r_2, \dots, r_m) = r \in \mathbb{R} \setminus \{0\}$. and $p(r_1, r_2, \dots, r_{i-1}, r_i^*, r_{i+1}, \dots, r_m) \notin \mathbb{R}$. Assume that $p(r_1, r_2, \dots, r_{i-1}, r_i^*, r_{i+1}, \dots, r_m) = a + v$ for $a \in \mathbb{R}$ and $v \in V$. Then $v \neq 0$. If $a = c \cdot p(r_1, r_2, \dots, r_m)$ we can take $\tilde{r}_i = r_i^* - cr_i$, and

$$p(r_1, r_2, \dots, r_{i-1}, \tilde{r}_i, r_{i+1}, \dots, r_m) = v \in V \setminus \{0\}.$$

Note that for arbitrary real numbers x and y we have an element

$$p(r_1, r_2, \dots, r_{i-1}, xr_i + yr_i\tilde{r}_i, r_{i+1}, \dots, r_m) = xr + yv.$$

Consider an arbitrary quaternion $a + bi + cj + dk$, where $a, b, c, d \in \mathbb{R}$. Let us take an x such that $xr = a$. By Lemma 2.32, V is an irreducible self-similar cone, thus there exist $h \in \mathbb{H}$ and $y \in \mathbb{R}$ such that $yhvh^{-1} = bi + cj + dk$. Hence,

$$\begin{aligned} p(hr_1h^{-1}, hr_2h^{-1}, \dots, hr_{i-1}h^{-1}, h(xr_i + y\tilde{r}_i)h^{-1}, hr_{i+1}h^{-1}, \dots, hr_mh^{-1}) \\ = h(xr + yv)h^{-1} = xr + yhvh^{-1} = a + bi + cj + dk. \end{aligned}$$

Therefore, $\text{Im } p = \mathbb{H}$. ■

2.1.6 Semi-homogeneous polynomials evaluated on \mathbb{H}

Now we can consider the semi-homogeneous case and present the proof of Theorem 1.26:

Proof of Theorem 1.26. There exists an isomorphism $\Phi: \mathbb{H}(\mathbb{R}) \otimes \mathbb{C}_{\mathbb{R}} \rightarrow M_2(\mathbb{C})_{\mathbb{R}}$ such that for any $q = a + bi + cj + dk \in \mathbb{H}$ and any $z \in \mathbb{C}$

$$\Phi(q \otimes z) = z \cdot \begin{bmatrix} a + bi & c + di \\ -c + di & a - bi \end{bmatrix}.$$

Note that $\Phi(\mathbb{H} \otimes 1_{\mathbb{C}})$, i.e., the set of matrices

$$\begin{bmatrix} a + bi & c + di \\ -c + di & a - bi \end{bmatrix}$$

is Zariski dense in $M_2(\mathbb{C})$.

Note that there exist numbers w_1, \dots, w_m and $d \neq 0$ such that for any $c \in \mathbb{R}$ $p(c^{w_1}x_1, c^{w_2}x_2, \dots, c^{w_m}x_m) = c^d p(x_1, \dots, x_m)$. Therefore $\text{Im } p$ must be a cone with respect to positive real multipliers, i.e., $\lambda x \in \text{Im } p$ for any $x \in \text{Im } p$ and any $\lambda > 0$.

By Theorem 2.3, the Zariski closure of the image of polynomial evaluated on 2×2 matrices with complex entries must be either $\{0\}$, or \mathbb{C} , or $\text{sl}_2(\mathbb{C})$ or $M_2(\mathbb{C})$.

In the first case p is PI, and its image is $\{0\}$. In the second case p is a central polynomial, therefore the image of p being a cone with respect to positive multipliers must be either \mathbb{R} , or $\mathbb{R}_{\geq 0}$, or $\mathbb{R}_{\leq 0}$. In the third case, p takes only pure quaternion values. According to Lemma 2.32 we know that V is an irreducible self-similar cone (up to real multipliers). Hence $\text{Im } p$ is a self-similar cone up to positive real multipliers. Therefore $\text{Im } p \cup (-\text{Im } p) = V$, i.e., for any pure quaternion $v \in V$, either v , or $-v$ belongs to $\text{Im } p$. Without loss of generality, assume that $i \in \text{Im } p$. Hence $ji j^{-1} \in \text{Im } p$, and $ji j^{-1} = -i$. Hence for any $c \in \mathbb{R}$, $ci \in V$, and thus any element $v \in V$ is conjugate to some ci , as we showed in the proof of Lemma 2.32. Therefore, $\text{Im } p = V$. In the fourth case an image of p is a Zariski dense subset of \mathbb{H} . ■

Let us show some interesting examples of multilinear and homogeneous polynomials:

Example 2.33.

- (i) According to Amitsur–Levitsky theorem s_4 evaluated on 2×2 matrices is a PI, and thus is PI on quaternions, as is its multilinearization.
- (ii) The Lie bracket $p(x, y) = xy - yx$ is a Lie polynomial and its image is V .
- (iii) The polynomial $[x_1, x_2][x_3, x_4] + [x_3, x_4][x_1, x_2]$, the multilinearization of $[x_1, x_2]^2$, evaluated on \mathbb{H} , is a multilinear central polynomial.

- (iv) An example where $\text{Im } p = \mathbb{H}(\mathbb{R})$ can be constructed trivially.
- (v) The central polynomial $p(x, y) = [x, y]^4$ provides an example of complete homogeneous polynomial evaluated on 2×2 matrices with real entries and taking only positive central values. I took a square of p . Note that evaluated on quaternions we do not need to take a square. This polynomial takes only non-positive values: indeed, $(ai + bj + ck)^2 = -a^2 - b^2 - c^2$, and $\text{Im } p$ is $\mathbb{R}_{\leq 0}$. Hence, $-p(x, y) = -[x, y]^2$ is an example of the polynomial with image set $\mathbb{R}_{\geq 0}$.
- (vi) The polynomial $p(x, y) = [x, y]^2 + [x^2, y^2]$ is the sum of two polynomials, the image of the first term is $\mathbb{R}_{\leq 0}$, and the second one has image V , thus $\text{Im } p$ is the set of quaternions with non-positive real part. Of course $-p$ has the opposite image: the set of quaternions with non-negative real part.
- (vii) In Example 2.8(i), the polynomial $g(x_1, x_2) = [x_1, x_2]^2$ has the property that $g(A, B) = 0$ whenever A is scalar, but g can take on a non-zero value whenever A is non-scalar. Thus, $g(x_1, x_2)x_1$ takes all values except scalars.
- (viii) Any quaternion $q = a + v$, $a \in \mathbb{R}$, $v \in V$ can be considered as a 2×2 matrix with complex entries, whose eigenvalues are $\lambda_{1,2} = a \pm ni$, where $n = \|v\|$ is the norm of the vector part of the quaternion. In particular $\text{tr } q = 2a = 2 \text{Re } q$ and $\det q = a^2 + n^2 = \|q\|^2$. In Example 2.8(ii) we provided an example of the polynomial taking all possible values except for those where the ratio of eigenvalues belongs to some set S . However here we should have a polynomial with real coefficients. Nevertheless this is possible: let S be any finite subset of S^2 – a unit circle on the complex plane. There exists a completely homogeneous polynomial p such that $\text{Im } p$ is the set of all quaternions except the quaternions with ratio of eigenvalues from S . Any $c \in S$ can be written in the form $c = \frac{a+bi}{a-bi}$ for some $a, b \in \mathbb{R}$. The construction of the polynomial is as follows. Consider

$$f(x) = x \cdot \prod_{c \in S} ((a+bi)\lambda_1 - (a-bi)\lambda_2)((a+bi)\lambda_2 - (a-bi)\lambda_1),$$

where $\lambda_{1,2}$ are eigenvalues of x and $c = \frac{a+bi}{a-bi}$. For each c the product $(a+bi)\lambda_1 - (a-bi)\lambda_2)((a+bi)\lambda_2 - (a-bi)\lambda_1) = -a^2(\lambda_1 - \lambda_2)^2 - b^2(\lambda_1 + \lambda_2)^2$ is a polynomial with real coefficients in $\text{tr } x$ and $\text{tr } x^2$. Thus $f(x)$ is a polynomial with traces, and by [92, Theorem 1.4.12], one can rewrite each trace in f as a fraction of multilinear central polynomials.

After that we multiply the expression by the product of all the denominators, which we can take to have value 1. We obtain a completely homogeneous polynomial p which image is the cone under $\text{Im } f$ and thus equals $\text{Im } f$. The image of p is the set of all quaternions with ratios of eigenvalues not belonging to S .

2.2 The Deligne trick

One of the strongest tools we use in our investigation is a famous trick of Deligne, given below in Remark 2.42. We use it when we give a proof by contradiction; in some cases we obtain an element of $\overline{\text{UD}}$ whose characteristic polynomial can be decomposed, and this contradicts to Amitsur's theorem. Here are some examples of when we use it.

2.2.1 Evaluations of word maps on matrix groups

It is important to investigate possible images of word maps on matrix groups, and the famous Deligne trick technique can be used in this area. A word map $w(x_1, \dots, x_m)$ is called dominant if its image is Zariski dense. A famous theorem of Borel states:

Theorem 2.34. *Any nontrivial word map w evaluated on $\mathrm{SL}_n(K)$ is dominant.*

Proof. Let us use induction with respect to n . The case $n = 1$ is trivial, since $\mathrm{SL}_1(K)$ is a trivial group, and any word map (including the trivial one) is dominant. Assume that Borel's theorem holds for some n and let us show that it holds for $n + 1$. Note that $\mathrm{SL}_n \leq \mathrm{SL}_{n+1}$, because one can take $\varphi: \mathrm{SL}_n \rightarrow \mathrm{SL}_{n+1}$ defined in the following way: for any matrix $M \in \mathrm{SL}_n$ we will take an $(n + 1) \times (n + 1)$ matrix from SL_{n+1} , having entries of M in the first n rows and columns, zeros on the last row and on the last column except an entry $(n + 1, n + 1)$ having 1 there. Therefore the Zariski closure of the image of w evaluated on SL_{n+1} , according to the induction assumption, is at least $\mathrm{Im} \varphi$. Note that $\mathrm{Im} w$ as well as its Zariski closure is closed under conjugations, therefore it will include all matrices from SL_{n+1} having at least one eigenvalue equal to 1. Let us show that this set cannot be Zariski closure of $\mathrm{Im} w$. Assume that it is. Note that w can be considered as an element of $\widetilde{\mathrm{UD}}$, and if 1 is its eigenvalue in any evaluation, then its characteristic polynomial has a divisor $\lambda - 1$, which is impossible, since $\chi_w(w) = (w - 1)f(w) = 0$ gives us that either $w = 1$ (which is impossible if w is not trivial), or $f(M) = 0$ for all M being values of the map w . Note that $\deg f = n$ and this is impossible since it contradicts Jordan's theorem. Here we use the Deligne trick. ■

If K is an algebraically closed field and $\mathrm{Char} K = 0$ then there is a working hypothesis that any nontrivial word map evaluated on $\mathrm{PSL}_n(K)$ must be surjective; here we will consider it for the case $n = 2$:

Conjecture 2.35. *If the field K is algebraically closed of characteristic 0, then the image of any nontrivial group word $w(x_1, \dots, x_m)$ on the projective linear group $\mathrm{PSL}_2(K)$ is $\mathrm{PSL}_2(K)$.*

Remark 2.36. Note that if one takes the group SL_2 instead of PSL_2 , Conjecture 2.35 fails, since the matrix $-I + e_{12}$ does not belong to the image of the word map $w = x^2$.

Example 2.37. When $\mathrm{Char} K = p > 0$, the image of the word map $w(x) = x^p$ evaluated on $\mathrm{PSL}_2(K)$ is not $\mathrm{PSL}_2(K)$. Indeed, otherwise the matrix $I + e_{12}$ could be written as x^p for $x \in \mathrm{PSL}_2(K)$. If the eigenvalues of x are equal, then $x = I + n$ where n is nilpotent. Therefore $x^p = (I + n)^p = I + pn = I$. If the eigenvalues of x are not equal, then x is diagonalizable and therefore x^p is also diagonalizable, a contradiction.

Lemma 2.38 (Liebeck, Nikolov, Shalev, cf. also [10, 44]). *$\mathrm{Im} w$ contains all matrices from $\mathrm{PSL}_2(K)$ which are not unipotent.*

Proof. According to [16] the image of the word map w must be Zariski dense in $\mathrm{SL}_2(K)$. Therefore the image of $\mathrm{tr} w$ must be Zariski dense in K . Note that $\mathrm{tr} w$ is a homogeneous rational function and K is algebraically closed. Hence, $\mathrm{Im}(\mathrm{tr} w) = K$. For any $\lambda \neq \pm 1$ any matrix with eigenvalues λ and λ^{-1} belongs to the image of w since there is a matrix with trace $\lambda + \lambda^{-1}$ in $\mathrm{Im} w$ and any two matrices from SL_2 with equal trace (except trace ± 2) are similar. ■

However the question of whether one of the matrices $(I + e_{12})$ or $(-I - e_{12})$ (which are equal in PSL_2) belongs to the image of w remains open.

The following Lie-algebraic counterpart of Borel's theorem holds:

Theorem 2.39 ([8]). *Let L be a split semisimple Lie algebra, k a field. Suppose that a Lie polynomial $w(x_1, \dots, x_n)$ is not an identity of the Lie algebra $\mathfrak{sl}_2(k)$. Then the image of $w: L^n \rightarrow L$ is Zariski dense.*

The proof of this Theorem uses Deligne trick.

Remark 2.40. At the best of our knowledge, the Deligne trick first appeared in the paper [28]. The Borel paper [16] was published in the same issue.

2.2.2 Evaluations of semihomogeneous polynomials on matrix rings

Proof of Theorem 2.3. Assume that there are matrices $p(x_1, \dots, x_m)$ and $p(y_1, \dots, y_m)$ with different ratios of eigenvalues in the image of p . Consider the polynomial matrix $f(t) = p(tx_1 + (1-t)y_1, tx_2 + (1-t)y_2, \dots, tx_m + (1-t)y_m)$, and $\Pi \circ f$ where Π is defined in Remark 2.5. Write this nonconstant rational function $\frac{\text{tr}^2 f}{\det f}$ in lowest terms as $\frac{A(t)}{B(t)}$, where $A(t), B(t)$ are polynomials of degree $\leq 2 \deg p$ in the numerator and denominator.

An element $c \in K$ is in $\text{Im}(\Pi \circ f)$ iff there exists t such that $A - cB = 0$ (If for some t^* $A(t^*) - cB(t^*) = 0$, then t^* would be a common root of A and B). Let $d_c = \deg(A - cB)$. Then $d_c \leq \max(\deg A, \deg B) \leq 2 \deg p$, and $d_c = \max(\deg A, \deg B)$ for almost all c . Hence, the polynomial $A - cB$ is not constant and thus there is a root. Thus the image of $\frac{A(t)}{B(t)}$ is Zariski dense, implying the image of $\frac{\text{tr}^2 f}{\det f}$ is Zariski dense.

Hence, we may assume that $\text{Im } p$ consists only of matrices having a fixed ratio r of eigenvalues. If $r \neq \pm 1$, the eigenvalues λ_1 and λ_2 are linear functions of $\text{tr } p(x_1, \dots, x_m)$. Hence λ_1 and λ_2 are the elements of the algebra of generic matrices with traces, which is a domain by Proposition 1.31. But the two nonzero elements $p - \lambda_1 I$ and $p - \lambda_2 I$ have product zero, a contradiction. Here we use the Deligne trick.

We conclude that $r = \pm 1$. First assume $r = 1$. If $\text{Char } K \neq 2$, then p is a PI, by Lemma 1.32. If $\text{Char } K = 2$ then the image is either $\text{sl}_2(K)$ or \hat{K} , by Example 2.2(v).

Thus, we may assume $r = -1$ and $\text{Char } K \neq 2$. Hence, $\text{Im } p$ consists only of matrices with $\lambda_1 = -\lambda_2$. By Lemma 1.32, there is a non-nilpotent matrix in the image of p . Hence, by Example 2.2(v), $\text{Im } p$ is either \hat{K} or strictly contains it and is all of $\text{sl}_2(K)$. ■

2.2.3 Evaluations of multilinear polynomials on matrix rings

Recall Lemma 2.14: *If p is a multilinear polynomial evaluated on the matrix ring $M_2(K)$ (where K is a quadratically closed field of characteristic not 2), then $\text{Im } p$ is either $\{0\}$, K , sl_2 , or $M_2(K)$.*

Remark 2.41. Since the details are rather technical, we start by sketching the proof. We assume that $\text{Im } p = M_2(K) \setminus \tilde{K}$. The linear change of the variable in position i gives us the line $A + tB$ in the image, where $A = p(x_1, \dots, x_m)$ and $B = p(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_m)$. Take the function that maps t to $f(t) = (\lambda_1 - \lambda_2)^2$, where λ_i are the eigenvalues of $A + tB$. Evidently

$$f(t) = (\lambda_1 - \lambda_2)^2 = (\lambda_1 + \lambda_2)^2 - 4\lambda_1\lambda_2 = (\text{tr}(A + tB))^2 - 4 \det(A + tB),$$

so our function f is a polynomial of $\deg \leq 2$ evaluated on entries of $A + tB$, and thus is a polynomial in t .

There are three possibilities: Either $\deg_t f \leq 1$, or f is the square of another polynomial, or f vanishes at two different values of t (say, t_1 and t_2). (Note that here we use that the field is quadratically closed). This polynomial f vanishes if and only if the two eigenvalues of $A + tB$ are equal, and this happens in two cases (according to Lemma 2.12): $A + tB$ is scalar or $A + tB$ is nilpotent. Thus either both $A + t_i B$ are scalar, or $A + t_1 B$ is scalar and $A + t_2 B$ is nilpotent, or both $A + t_i B$ are nilpotent. In the first instance A and B are scalars, which is impossible. The second case instance that the matrix $A + \frac{t_1+t_2}{2} B \in \tilde{K}$, which is also impossible. The third instance implies that $\text{tr } A = \text{tr } B = 0$ which we claim is also impossible. If $\deg_t f \leq 1$, then for large t the difference $\lambda_1 - \lambda_2$ of the eigenvalues of $A + tB$, is much less than t , so the difference between eigenvalues of B must be 0, a contradiction.

It follows that $f(t) = (\lambda_1 - \lambda_2)^2$ is the square of a polynomial (with respect to t). Thus $\lambda_1 - \lambda_2 = a + tb$, where a and b are some functions of the entries of the matrices x_1, \dots, x_m, y_i . Note that a is the difference of eigenvalues of A and b is the difference of eigenvalues of B , Thus

$$a(x_1, \dots, x_m, y_i) = b(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_m, x_i). \quad (2.2)$$

Note that $(\lambda_1 - \lambda_2)^2 = a^2 + 2abt + b^2t^2$ which means that a^2 , b^2 and ab are polynomials (note that here we use $\text{Char } K \neq 2$). Thus, $\frac{a}{b} = \frac{a^2}{ab}$ is a rational function. Therefore there are polynomials p_1 , p_2 and q such that $a = p_1\sqrt{q}$ and $b = p_2\sqrt{q}$. Without loss of generality, q does not have square divisors. By (2.2) we have that q does not depend on x_i and y_i . Now consider the change of other variables. The function a is the difference of eigenvalues of $A = p(x_1, \dots, x_m)$ so it remains unchanged. Thus q does not depend on other variables also. That is why $\lambda_1 \pm \lambda_2$ are two polynomials and hence λ_i are polynomials. One concludes with the last paragraph of the proof.

Proof of Lemma 2.14. In view of Lemma 2.12 it suffices to prove that the image of p cannot be $M_2(K) \setminus \tilde{K}$. Assume that the image of p is $M_2(K) \setminus \tilde{K}$. Consider for each variable x_i the line $x_i + ty_i$, $t \in K$. Then $p(x_1, \dots, x_{i-1}, x_i + ty_i, x_{i+1}, \dots, x_m)$ is the line $A + tB$, where $p(x_1, \dots, x_m) = A$ and $p(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_m) = B$. Thus $A + tB \notin \tilde{K}$ for any t . Since B is diagonalizable, we can choose our matrix units $e_{i,j}$ such that B is diagonal. Therefore

$$B = \lambda_B I + \begin{pmatrix} c & 0 \\ 0 & -c \end{pmatrix}, \quad A = \lambda_A I + \begin{pmatrix} x & y \\ z & -x \end{pmatrix}.$$

Hence

$$A + tB = (\lambda_A + t\lambda_B)I + \begin{pmatrix} x + tc & y \\ z & -x - tc \end{pmatrix}.$$

The matrix $\begin{pmatrix} x+tc & y \\ z & -x-tc \end{pmatrix}$ is nilpotent if and only if $(x + tc)^2 + yz = 0$, which has the solution $t_{1,2} = \frac{1}{c}(-x \pm \sqrt{-yz})$. Thus, when $yz \neq 0$, $\pi(A + t_j B)$ will be nilpotent for $j = 1, 2$, where $\pi(X) = X - \frac{1}{2} \text{tr } X$. However $\text{tr}(A + t_j B)$ is nonzero for one of these values of t_j , implying $A + t_j B \in \tilde{K}$, a contradiction.

Thus, we must have $yz = 0$. Without loss of generality we can assume that $z = 0$. Any matrix M of the type $qI + \begin{pmatrix} w & h \\ 0 & -w \end{pmatrix}$ satisfies $\det M = q^2 - w^2$ and $q = \frac{1}{2} \text{tr } M$. Thus $x = \sqrt{\frac{1}{4}(\text{tr } A)^2 - \det A}$ and $c = \sqrt{\frac{1}{4}(\text{tr } B)^2 - \det B}$. Consider the matrix

$$P_i = cA - xB = p(x_1, \dots, x_{i-1}, cx_i - xy_i, x_{i+1}, \dots, x_m),$$

which must be scalar or nilpotent. It can be written explicitly algebraically in terms of the entries of x_i and y_i . Also, $P_i = (c\lambda_B - x\lambda_A)I + (cy)e_{12}$, where e_{12} is the matrix unit. There are two cases. If $y = 0$ then the line $A + tB$ includes a scalar matrix, and if $y \neq 0$ then $(c\lambda_B - x\lambda_A) = 0$ and all matrices on the line $A + tB$ have the same ratio of eigenvalues.

Let $S_1 = \{i: P_i \in K\}$ and $S_2 = \{i: P_i \in \mathfrak{sl}_2(K)\}$. Without loss of generality we can assume for some $k \leq m$ that $S_1 = \{1, 2, \dots, k\}$ and $\{k + 1, \dots, m\}$. The four entries of $p(x_1, \dots, x_m)$ are

$$p_{ij}(x_{1,(1,1)}, x_{1,(1,2)}, x_{1,(2,1)}, x_{1,(2,2)}, \dots, x_{m,(2,2)}),$$

polynomials in the entries of x_i . Consider the scalar function

$$f_1(x_1, \dots, x_m) = \frac{\frac{1}{2} \text{tr } p(x_1, \dots, x_m)}{R(x_1, \dots, x_m)},$$

where $R(x_1, \dots, x_m) = \sqrt{\frac{1}{4} \text{tr}^2 p(x_1, \dots, x_m) - \det p(x_1, \dots, x_m)}$. This function is defined everywhere except for those (x_1, \dots, x_m) for which $p(x_1, \dots, x_m)$ is a matrix with equal eigenvalues, because R is the half-difference of eigenvalues. The function $f_1(x_1, \dots, x_m)$ does not depend on

x_{k+1}, \dots, x_m because for any $i \geq k+1$, substituting y_i instead of x_i does not change the ratio of eigenvalues of $p(x_1, \dots, x_m)$. Consider the matrix function

$$f_2(x_1, \dots, x_m) = \frac{p(x_1, \dots, x_m) - \frac{1}{2} \operatorname{tr} p(x_1, \dots, x_m)}{R(x_1, \dots, x_m)}.$$

This function is also defined everywhere except for those (x_1, \dots, x_m) such that the eigenvalues of $p(x_1, \dots, x_m)$ are equal. The function $f_2(x_1, \dots, x_m)$ does not depend on x_i , $i \leq k$, because for any $i \leq k$ substituting y_i instead of x_i does not change the basis in which $p(x_1, \dots, x_m)$ is diagonal. R^2 is a polynomial:

$$R^2 = \frac{1}{4} \operatorname{tr}^2 p(x_1, \dots, x_m) - \det p(x_1, \dots, x_m).$$

Write $R^2 = r_1 r_2 r_3$ where r_1 is the product of all the irreducible factors in which only x_1, \dots, x_k occur, r_2 is the product of all the irreducible factors in which only x_{k+1}, \dots, x_m occur, r_3 is the product of the other irreducible factors. We have that

$$\frac{\operatorname{tr}^2 p(x_1, \dots, x_m)}{r_1(x_1, \dots, x_m) r_2(x_1, \dots, x_m) r_3(x_1, \dots, x_m)} = f_1^2(x_1, \dots, x_m)$$

does not depend on x_{k+1}, \dots, x_m . Therefore if $\operatorname{tr}^2 p = q_1 q_2 q_3$ (again in q_1 only x_1, \dots, x_k occur, in q_2 only x_{k+1}, \dots, x_m occur and q_3 is all the rest) then $\frac{r_1 r_2 r_3}{q_1 q_2 q_3}$ does not depend on x_{k+1}, \dots, x_m . Hence $r_2 = q_2$ and $r_3 = q_3$ (up to scalar factors). As soon as $q_1 q_2 q_3$ is a square of a polynomial all q_i are squares therefore r_2 and r_3 are squares. Now consider the function $\frac{p_{12}^2}{R^2}$. This is the square of the (1, 2)-entry in the matrix function f_2 , so it does not depend on x_1, \dots, x_k . Writing $p_{12}^2 = q_1 q_2 q_3$ (where, again, only x_1, \dots, x_k occur in q_1 , only x_{k+1}, \dots, x_m occur in q_2 and q_3 is comprised of all the rest), then all the q_i are squares and $q_1 = r_1$, implying r_1 is square. Thus the polynomial $r_1 r_2 r_3 = R^2$ is the square of a polynomial. Therefore R is a polynomial. We conclude that $\lambda_1 - \lambda_2 = 2R$ is a polynomial (where we recall that λ_1 and λ_2 are the eigenvalues of $p(x_1, \dots, x_m)$). $\lambda_1 + \lambda_2 = \operatorname{tr}(p)$ is also a polynomial and hence λ_i are polynomials, which obviously are invariant under conjugation since any conjugation is the square of some other conjugation). Hence, λ_i are the polynomials of traces, by Donkin's theorem quoted above. Now consider the polynomials $(p - \lambda_1 I)$ and $(p - \lambda_2 I)$, which are elements of the algebra of free matrices with traces, which we noted above is a domain. Both are not zero but their product is zero, a contradiction. ■

Remark 2.42. This trick (when we use that the characteristic polynomial of a non-central element of $\widetilde{\text{UD}}$ cannot be decomposed) is called the *Deligne trick*.

2.3 Some important questions and open problems regarding evaluations of polynomials on low rank algebras

In this section we formulate some problems that remain open.

The first and one of the most important problems in this area is the investigation of the possible image sets of multilinear polynomials on matrix algebras. One of the main and the most perspective conjectures answering this question is the L'vov–Kaplansky conjecture. However it remains being an open problem. For the case of 2×2 matrices we have Theorem 1.15, nevertheless the case when $\operatorname{sl}_2 \subsetneq \operatorname{Im} p$ should be investigated, not only for $K = \mathbb{R}$.

Problem 2.43. Give a full classification of all possible image sets of multilinear polynomials evaluated on $M_2(K)$ for arbitrary field K in the case when $\operatorname{sl}_2 \subsetneq \operatorname{Im} p$.

Another important problem is to investigate possible images of semihomogeneous polynomials evaluated on $M_2(K)$. We have a good result when K is quadratically closed (see Theorem 1.13), however our result for $K = \mathbb{R}$ does not settle the problem. For \mathbb{R} the Zariski topology is not so natural, and it would be much more important to investigate possible images with respect to the standard topology. This question remains open. The question about classification of such image sets for the arbitrary field remains being open as well. The same question can be asked for the quaternion algebra. We have Theorem 1.26, however the problem of classification of all possible Zariski dense images of semi-homogeneous polynomials evaluated on \mathbb{H} remains being open.

Problem 2.44. *Give a classification of all possible image sets of semihomogeneous polynomials evaluated on $M_2(K)$ for arbitrary K , or at least for $K = \mathbb{R}$ with respect to the standard topology. Give a classification of all possible images of semihomogeneous polynomials evaluated on \mathbb{H} with respect to the standard topology.*

The same question can be asked for arbitrary (non-homogeneous) polynomial. Let us provide an important example:

Example 2.45. Let $h(x)$ be arbitrary polynomial in one variable, and consider a noncommutative polynomial $h([x, y])$ evaluated on $M_2(K)$. Note that $[x, y]$ is trace zero matrix, so its eigenvalues are $\pm\lambda$ for some λ . Thus, if $[x, y]$ is not nilpotent then $[x, y]$ is similar to the matrix $\lambda e_{11} - \lambda e_{22}$. Hence, $h([x, y])$ is similar to the matrix $h(\lambda)e_{11} + h(-\lambda)e_{22}$, and its eigenvalues are $h(\pm\lambda)$. Therefore, $\text{Im } h([x, y])$ is the set of all matrices having pairs of eigenvalues $h(\pm\lambda)$.

The question whether all possible images of noncommutative polynomials evaluated on $M_2(K)$ can be only $\{0\}$, K , $\text{sl}_2(K)$, $M_2(K)$ or the set of all matrices having pairs of eigenvalues $h(\pm\lambda)$ for some polynomial h is an open problem. The same question can be asked for the evaluations of non homogeneous polynomials on the quaternion algebra.

Problem 2.46. *Give a classification of all possible image sets of non homogeneous (arbitrary) polynomials on $M_2(K)$ (and \mathbb{H}), where K is an arbitrary field (or for some partial cases such as $K = \mathbb{R}$, $K = \mathbb{C}$) K being any quadratically closed field.*

Another important algebra is the algebra of Cayley numbers. This algebra is nonassociative, and therefore polynomials evaluated on it are also nonassociative. However, it is close to associative, and part of our tools can work. The problem is to give a classification of possible image sets of nonassociative multilinear polynomials evaluated on the algebra of Cayley numbers and to check whether all possible images are vector spaces.

Problem 2.47. *Give a classification of all possible evaluations of multilinear polynomials on the Cayley numbers algebra, the next step would be to consider the same question for semihomogeneous and for arbitrary polynomials.*

Question 2.48. *Is it true for a simple 3-dimensional Lie algebra that it is possible not only to provide a classification of possible image sets of polynomials (see Theorem 1.23) but also for any given set S to describe the set of all polynomials whose image sets are S ?*

Question 2.49 (see Conjecture 2.35). *If the field K is algebraically closed of characteristic 0, then must the image of any nontrivial group word $w(x_1, \dots, x_m)$ on the projective linear group $\text{PSL}_2(K)$ be $\text{PSL}_2(K)$?*

Malev and Pines [83] have established that the evaluations of multilinear polynomials on the rock-paper-scissors algebra is a vector space.

3 The case of matrices of rank 3

Now we turn specifically to the case $n = 3$ and the proof of Theorem 1.20.

Lemma 3.1. *We define functions $\omega_k: M_3(K) \rightarrow K$ as follows: Given a matrix a , let $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues of a , and denote*

$$\omega_k := \omega_k(a) = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq 3} \lambda_{i_1} \dots \lambda_{i_k}.$$

Let $p(x_1, \dots, x_m)$ be a semi-homogeneous, trace vanishing polynomial.

Consider the rational function $H(x_1, \dots, x_m) = \frac{\omega_2(p(x_1, \dots, x_m))^3}{\omega_3(p(x_1, \dots, x_m))^2}$ (taking values in $K \cup \{\infty\}$). If $\text{Im } H$ is dense in K , then $\text{Im } p$ is dense in sl_3 .

Proof. Note that $\omega_2(p)^3$ and $\omega_3(p)^2$ are semi-homogeneous. Thus, $\text{Im } H$ is dense in K iff the image of the pair $(\omega_2(p)^3, \omega_3(p)^2)$ is dense in K^2 . But since ω_2 and ω_3 are algebraically independent, so are $\omega_2(p)^3$ and $\omega_3(p)^2$, so we conclude that the image of the pair $(\omega_2(p)^3, \omega_3(p)^2)$ is dense in K^2 . Thus, the set of characteristic polynomials of evaluations of p is dense in the space of all possible characteristic polynomials of trace zero matrices. Therefore, the set of all triples $(\lambda_1, \lambda_2, -\lambda_1 - \lambda_2)$ of eigenvalues of matrices from $\text{Im } p$ is dense in the plane $x + y + z = 0$ defined in K^3 , implying that $\text{Im } p$ is dense in sl_3 . ■

3.1 3×3 matrices over a field with a primitive cube root of 1

Let K be an algebraically closed field. For $\text{char}(K) \neq 3$ we fix a primitive cube root $\varepsilon \neq 1$ of 1; when $\text{char}(K) = 3$ we take $\varepsilon = 1$.

The proof of Theorem 1.18 will be presented in Section 3.3.

Example 3.2. The element $[x, [y, x]x[y, x]^{-1}]$ of $\widetilde{\text{UD}}$ takes on only 3-scalar values (see [92, Theorem 3.2.21, p. 180]) and thus gives rise to a homogeneous polynomial taking on only 3-scalar values.

3.2 Multilinear trace vanishing polynomials

We can characterize the possible image sets of multilinear trace vanishing polynomials.

Lemma 3.3. *If p is a multilinear polynomial, not PI nor central, then there exists a collection of matrix units (e_1, e_2, \dots, e_m) such that $p(e_1, e_2, \dots, e_m)$ is a diagonal but not scalar matrix.*

Proof. According to Lemma 1.27, for any matrix units e_i the value $p(e_1, e_2, \dots, e_m)$ must be either $ce_{i,j}$ or some diagonal matrix. In addition, let us note that according to the Remark 1.3 the linear span of the image set $\langle p(M_3) \rangle$ (being a linear span of all evaluations of p on sets of matrix units) is one of four possible sets: either is either $\{0\}$, K , $\text{sl}_3(K)$, or $M_3(K)$. We assumed that p is not PI nor central, hence the linear span of the image set $\langle p(M_3) \rangle$ either is $\text{sl}_n(K)$, or $M_n(K)$. Thus, the linear span of all $p(e_1, e_2, \dots, e_m)$ for any matrix units e_i such that $p(e_1, e_2, \dots, e_m)$ is diagonal, includes all $\text{Diag}\{x, y, -x - y\}$. In particular there exists a collection of matrix units (e_1, e_2, \dots, e_m) such that $p(e_1, e_2, \dots, e_m)$ is a diagonal but not scalar matrix. ■

The proof of Theorem 1.19 will be presented in Section 3.3.

Remark 3.4. Assume that $\text{char}(K) = 3$ and p is a multilinear polynomial, which is neither PI nor central. Then, by Lemma 3.3 there exists a collection of matrix units e_i such that

$$p(e_1, \dots, e_m) = \text{Diag}\{\alpha, \beta, \gamma\}$$

is diagonal but not scalar. Without loss of generality, $\alpha \neq \beta$. Hence $p^3(e_1, \dots, e_m) = \text{Diag}\{\alpha^3, \beta^3, \gamma^3\}$ and $\alpha^3 \neq \beta^3$ because $\text{char}(K) = 3$. Therefore p is not 3-scalar.

Theorem 3.5. *If p is a multilinear polynomial such that Imp does not satisfy the equation $\gamma\omega_1(p)^2 = \omega_2(p)$ for $\gamma = 0$ or $\gamma = \frac{1}{4}$, then Imp contains a matrix with two equal eigenvalues that is not diagonalizable and of determinant not zero. If Imp does not satisfy any equation of the form $\gamma\omega_1(p)^2 = \omega_2(p)$ for any γ , then the set of non-diagonalizable matrices of Imp is Zariski dense in the set of all non-diagonalizable matrices, and Imp is dense.*

Proof. If not, then by Lemma 1.34 there is at least one variable (say, x_1) such that $a = p(x_1, x_2, \dots, x_m)$ does not commute with $b = p(\tilde{x}_1, x_2, \dots, x_m)$. Consider the matrix $a + tb = p(x_1 + t\tilde{x}_1, x_2, \dots, x_m)$, viewed as a polynomial in t .

Recall that the discriminant of a 3×3 matrix with eigenvalues $\lambda_1, \lambda_2, \lambda_3$ is defined as $\prod_{1 \leq i < j \leq 3} (\lambda_i - \lambda_j)^2$. Thus, the discriminant of $a + tb$ is a polynomial $f(t)$ of degree 6. If $f(t)$ has only one root t_0 , then this root is defined in terms of the entries of $\tilde{x}_1, x_1, x_2, \dots, x_m$, and invariant under the action of the symmetric group, and thus is in Amitsur's division algebra \widetilde{UD} . By Lemma 1.49, $a + t_0b$ is scalar, and the uniqueness of t_0 implies that a and b are scalar, contrary to assumption.

Thus, $f(t)$ has at least two roots – say, $t_1 \neq t_2$, and the matrices $a + t_1b$ and $a + t_2b$ each must have multiple eigenvalues. If both of these matrices are diagonalizable, then each of $a + t_ib$ have a 2-dimensional plane of eigenvectors. Therefore we have two 2-dimensional planes in 3-dimensional linear space, which must intersect. Hence there is a common eigenvector of both $a + t_ib$ and this is a common eigenvector of a and b . If a and b have a common eigenspace of dimension 1 or 2, then there is at least one eigenvector (and thus eigenvalue) of a that is uniquely defined, implying $a \in \widetilde{UD}$ by Remark 1.42, contradicting Lemma 1.49. If a and b have a common eigenspace of dimension 3, then a and b commute, a contradiction.

We claim that there cannot be a diagonalizable matrix with equal eigenvalues on the line $a + tb$. Indeed, if there were such a matrix, then either it would be unique (and thus an element of \widetilde{UD} , which cannot happen), or there would be at least two such diagonalizable matrices, which also cannot happen, as shown above.

Assume that all matrices on the line $a + tb$ of discriminant zero have determinant zero. Then either all of them are of the type $\text{Diag}\{\lambda, \lambda, 0\} + e_{12}$ or all of them are of the type $\text{Diag}\{0, 0, \mu\} + e_{12}$. (Indeed, there are three roots of the determinant equation $\det(a + tb) = 0$, which are pairwise distinct, and all of them give a matrix with two equal eigenvalues, all belonging to one of these types, since otherwise one eigenvalue is uniquely defined and thus yields an element of \widetilde{UD} , which cannot happen.)

In the first case, all three roots of the determinant equation $\det(a + tb) = 0$ satisfy the equation $(\omega_1(a + tb))^2 = 4\omega_2(a + tb)$. Hence, we have three pairwise distinct roots of the polynomial of maximal degree 2, which can occur only if the polynomial is identically zero. It follows that also $(\omega_1(a))^2 - 4\omega_2(a) = 0$, so $(\omega_1(p))^2 - 4\omega_2(p) = 0$ is identically zero, which by hypothesis cannot happen.

In the second case we have the analogous situation, but $\omega_2(p)$ will be identically zero, a contradiction.

Thus on the line $a + tb$ we have at least one matrix of the type $\text{Diag}\{\lambda, \lambda, \mu\} + e_{12}$ and $\lambda\mu \neq 0$. Consider the algebraic expression $\mu\lambda^{-1}$. If not constant, then it takes on almost all values, so assume that it is a constant δ . Then $\delta \neq -2$, since otherwise this matrix will be the unique matrix of trace 0 on the line $a + tb$ and thus an element of \widetilde{UD} , contrary to Lemmas 1.49 and 1.35. Consider the polynomial $q = p - \frac{\text{tr } p}{\delta + 2}$. At the same point t it takes on the value $\text{Diag}\{0, 0, (\delta - 1)\lambda\} + e_{12}$. Hence all three pairwise distinct roots of the equation $\det q(x_1 + t\tilde{x}_1, x_2, \dots, x_m) = 0$ will give us a matrix of the form $\text{Diag}\{0, 0, *\} + e_{12}$ (otherwise we have uniqueness and thus an element of \widetilde{UD}), contradicting Lemma 1.35. Therefore q satisfies an equation $\omega_2(q) = 0$. Hence, p satisfies an equation $\omega_1(p)^2 - c\omega_2(p) = 0$, for some constant c , a contradiction. Hence almost all non-diagonalizable matrices belong to the image of p , and they

are almost all matrices of discriminant 0 (a subvariety of $M_3(K)$ of codimension 1). By Amitsur's theorem, $\text{Im } p$ cannot be a subset of the discriminant surface. Thus, $\text{Im } p$ is dense in $M_3(K)$. ■

Remark 3.6. Note that if $\omega_1(p)$ is identically zero, and $\omega_2(p)$ is not identically zero, then $\text{Im } p$ contains a matrix similar to $\text{Diag}\{1, 1, -2\} + e_{12}$. Hence $\text{Im } p$ contains all diagonalizable trace zero matrices (perhaps with the exception of the diagonalizable matrices of discriminant 0, i.e., matrices similar to $\text{Diag}\{c, c, -2c\}$), all non-diagonalizable non-nilpotent trace zero matrices, and all matrices N for which $N^2 = 0$. Nilpotent matrices of order 3 also belong to the image of p , as we shall see in Lemma 3.8.

3.2.1 Multilinear trace vanishing polynomials over an algebraically closed field

Lemma 3.7. *A matrix is 3-scalar iff its eigenvalues are in $\{\gamma, \gamma\varepsilon, \gamma\varepsilon^2: \gamma \in K\}$, where $\gamma^3 \in K$ is its determinant. The variety V_3 of 3-scalar matrices has dimension 7.*

Proof. The first assertion is immediate since the characteristic polynomial is $x^3 - \gamma^3$. Hence V_3 is a variety. The second assertion follows since the invertible elements of V_3 are defined by two equations: $\text{tr}(x) = 0$ and $\text{tr}(x^{-1}) = 0$ and thus a V_3 is a variety of codimension 2. ■

Lemma 3.8. *Assume $\text{Char } K \neq 3$. If p is neither PI nor central, then the variety V_3 is contained in $\text{Im } p$.*

Proof. According to Lemma 1.27 there exist matrix units e_1, e_2, \dots, e_m such that $p(e_1, e_2, \dots, e_m) = e_{1,2}$. Consider the mapping χ described in the proof of Theorem 3.10 (see Section 3.4 for details). For any triples $T_i = (t_{1,i}, t_{2,i}, t_{3,i})$, let

$$f(T_1, T_2, \dots, T_m) = p(\dots, t_{1,i}e_i + t_{2,i}\chi(e_i) + t_{3,i}\chi^2(e_i), \dots).$$

$\text{Im } f$ (a subset of $\text{Im } p$) is a subset of the 3-dimensional linear space

$$L = \{\alpha e_{12} + \beta e_{23} + \gamma e_{31}, \alpha, \beta, \gamma \in K\}.$$

Since e_{12}, e_{23} and e_{31} belong to $\text{Im } f$, we see that $\text{Im } f$ is dense in L , and hence at least one matrix $a = \alpha e_{12} + \beta e_{23} + \gamma e_{31}$ for $\alpha\beta\gamma \neq 0$ belongs to $\text{Im } p$. Note that this matrix is 3-central. Thus the variety V_3 , excluding the nilpotent matrices, is contained in $\text{Im } p$. The nilpotent matrices of order 2 also belong to the image of p since they are similar to e_{12} .

Let us show that all nilpotent matrices of order 3 (i.e., matrices similar to $e_{12} + e_{23}$), also belong to $\text{Im } p$. We have the multilinear polynomial

$$f(T_1, T_2, \dots, T_m) = q(T_1, T_2, \dots, T_m)e_{12} + r(T_1, T_2, \dots, T_m)e_{23} + s(T_1, T_2, \dots, T_m)e_{31},$$

therefore q, r and s are three scalar multilinear polynomials. Assume there is no nilpotent matrix of order 3 in $\text{Im } p$. Then we have the following: if $q = 0$ then either $rs = 0$, if $r = 0$ then $sq = 0$, and if $s = 0$ then $qr = 0$. Assume q_1 is the greatest common divisor of q and r and $q_2 = \frac{q}{q_1}$. Note that both q_i are multilinear polynomials defined on disjoint sets of variables. If $q_1 = 0$ then $r = 0$ and if $q_2 = 0$ then $s = 0$. Since there are no repeated factors, $r = q_1 r'$ is a multiple of q_1 and $s = q_2 s'$ is a multiple of q_2 . The polynomial r' cannot have common divisors with q_2 , therefore if we consider any generic point (T_1, \dots, T_m) on the surface $r' = 0$ then $r(T_1, \dots, T_m) = 0$ and $q(T_1, \dots, T_m) \neq 0$. Hence $s(T_1, \dots, T_m) = 0$ for any generic (T_1, \dots, T_m) from the surface $r' = 0$. Therefore r' is the divisor of s . Remind both q_1 and q_2 are multilinear polynomials defined on disjoint subsets of $\{T_1, T_2, \dots, T_m\}$. Without loss of generality $q_1 = q_1(T_1, \dots, T_k)$, and $q_2 = q_2(T_{k+1}, \dots, T_m)$. Therefore $r' = r'(T_{k+1}, \dots, T_m)$ and it is divisor of s . Also remind $s = s'q_2$ so $q_2(T_{k+1}, \dots, T_m)$ is also divisor of s . Hence $r' = cq_2$ where c is constant. Thus $r = q_1 r' = cq_1 q_2 = cq$. However there exist (T_{k+1}, \dots, T_m) such that $q = 0$ and $r = 1$ (i.e., such that $f(T_{k+1}, \dots, T_m) = e_{23}$). A contradiction. ■

Remark 3.9. When $\text{Char } K = 3$, V_3 is the space of matrices with equal eigenvalues (including also scalar matrices). The same proof shows that all nilpotent matrices belong to the image of p , as well as all matrices similar to $cI + e_{12} + e_{23}$. But we do not know how to show that scalar matrices and matrices similar to $cI + e_{12}$ belong to the image of p .

Proof of Theorem 1.20. First assume that $\text{Char } K \neq 3$. According to Lemma 3.8 the variety V_3 is contained in $\text{Im } p$. Therefore $\text{Im } p$ is either the set of 3-scalar matrices, or some 8-dimensional variety (with 3-scalar subvariety), or is 9-dimensional (and thus dense).

It remains to classify the possible 8-dimensional images. Let us consider all matrices $p(e_1, \dots, e_m)$ where e_i are matrix units. If all such matrices have trace 0, then $\text{Im } p$ is dense in $\text{sl}_3(K)$, by Theorem 1.19. Therefore we may assume that at least one such matrix a has eigenvalues α, β and γ such that $\alpha + \beta + \gamma \neq 0$. By Theorem 3.10 we cannot have $\alpha + \beta + \gamma, \alpha + \beta\varepsilon + \gamma\varepsilon^2$ and $\alpha + \beta\varepsilon^2 + \gamma\varepsilon$ all nonzero. Hence a either is scalar, or a linear combination (with nonzero coefficients) of a scalar matrix and $\text{Diag } \{1, \varepsilon, \varepsilon^2\}$ (or with $\text{Diag } \{1, \varepsilon^2, \varepsilon\}$, without loss of generality – with $\text{Diag } \{1, \varepsilon, \varepsilon^2\}$). By Theorem 3.5, if $\text{Im } p$ is not dense, then p satisfies an equation of the type $(\text{tr}(p))^2 = \gamma \text{tr}(p^2)$ for some $\gamma \in K$. Therefore, if a scalar matrix belongs to $\text{Im } p$, then $\gamma = \frac{1}{3}$ and $\text{Im } p$ is the set of 3-scalar plus scalar matrices. If the matrix a is not scalar, then it is a linear combination of a scalar matrix and $\text{Diag } \{1, \varepsilon, \varepsilon^2\}$. Hence, by Remark 3.11, $\text{Im } p$ is also the set of 3-scalar plus scalar matrices. At any rate, we have shown that $\text{Im } p$ is either $\{0\}, K$, the set of 3-scalar matrices, the set of 3-scalar plus scalar matrices (matrices with eigenvalues $(\alpha + \beta, \alpha + \beta\varepsilon, \alpha + \beta\varepsilon^2)$), $\text{sl}_3(K)$ (perhaps lacking nilpotent matrices of order 3), or is dense in $M_3(K)$.

If $\text{Char } K = 3$, then by Remark 3.12 the multilinear polynomial p is either trace vanishing or $\text{Im } p$ is dense in $M_3(K)$. If p is trace vanishing, then by Theorem 1.19, $\text{Im } p$ is one of the following: $\{0\}$, the set of scalar matrices, the set of 3-scalar matrices, or for each triple $\lambda_1 + \lambda_2 + \lambda_3 = 0$ there exists a matrix $M \in \text{Im } p$ with eigenvalues λ_1, λ_2 and λ_3 . ■

3.3 Deligne trick for algebras of rank 3

Here we show proofs of two important theorems using the Deligne trick (for details see Section 2.2).

3.3.1 Evaluations of semihomogeneous polynomials on 3×3 matrices

Proof of Theorem 1.18. We define the functions $\omega_k: M_n(K) \rightarrow K$ as in Lemma 3.1, and consider the rational function $H = \frac{\omega_2(p(x_1, \dots, x_m))^3}{\omega_3(p(x_1, \dots, x_m))^2}$ (taking values in $K \cup \{\infty\}$).

If $\omega_2(p) = \omega_3(p) = 0$, then each evaluation of p is a nilpotent matrix, contradicting Amitsur's theorem. Thus, either $\text{Im } H$ is dense in K , or H must be constant.

If $\text{Im } H$ is dense in K , then $\text{Im } p$ is dense in sl_3 by Lemma 3.1.

So we may assume that H is a constant, i.e., $\alpha\omega_2^3(p) + \beta\omega_3^2(p) = 0$ for some $\alpha, \beta \in K$ not both 0. Fix generic matrices Y_1, \dots, Y_m . We claim that the eigenvalues $\lambda_1, \lambda_2, -\lambda_1 - \lambda_2$ of $q := p(Y_1, \dots, Y_m)$ are pairwise distinct. Otherwise either they are all equal, or two of them are equal and the third is not, each of which is impossible by Lemmas 1.49 and 1.35 since $q \in \widetilde{\text{UD}}$.

Let $\lambda'_1, \lambda'_2, -\lambda'_1 - \lambda'_2$ be the eigenvalues of another matrix $r \in \text{Im } p$. Thus we have the following:

$$\alpha\omega_2^3(r) + \beta\omega_3^2(r) = \alpha\omega_2^3(q) + \beta\omega_3^2(q) = 0.$$

Therefore we have homogeneous equations on the eigenvalues. Dividing by λ_2^6 and $\lambda_2'^6$ respectively, we have the same two polynomial equations of degree 6 on $\frac{\lambda_1}{\lambda_2}$ and $\frac{\lambda_1'}{\lambda_2'}$, yielding six possibilities for $\frac{\lambda_1'}{\lambda_2'}$. The six permutations of λ_1, λ_2 , and $\lambda_3 = -\lambda_1 - \lambda_2$ define six pairwise

different $\frac{\lambda_1^i}{\lambda_2^i}$ unless $(\lambda_1, \lambda_2, \lambda_3)$ is a permutation (multiplied by a scalar) of one of the following triples: $(1, 1, -2)$, $(1, -1, 0)$, $(1, \varepsilon, \varepsilon^2)$. The first instance is impossible since the eigenvalues must be pairwise distinct. Here we use the Deligne trick. The second instance give us an element of Amitsur's algebra $\widetilde{\text{UD}}$ with eigenvalue 0 and thus determinant 0, contradicting Amitsur's theorem. Here we use the Deligne trick again. In the third instance the polynomial p is 3-scalar. Thus, either p is 3-scalar, or each matrix from $\text{Im } p$ will have the same eigenvalues up to permutation and scalar multiple (which also holds when p is 3-scalar).

Assume that for some $i \in \{2, 3\}$ that $\text{tr } p^i$ is not identically zero. Then $\lambda_1^i, \lambda_2^i, \text{ and } \lambda_3^i$ are three linear functions on $\text{tr } p^i$. Hence we have the PI (polynomial identity) $(p^i - \lambda_1^i)(p^i - \lambda_2^i)(p^i - \lambda_3^i)$. Thus by Amitsur's theorem, one of the factors is a PI. Hence p^i is a scalar matrix. However $i \neq 2$ by Lemma 1.35. Hence $i = 3$, implying $\text{Im } p$ is the set of matrices with eigenvalues $\{(\gamma, \gamma\varepsilon, \gamma\varepsilon^2) : \gamma \in K\}$.

Thus, we may assume that p satisfies $\text{tr}(p^i) = 0$ for $i = 1, 2$ and 3. Now $\omega_1(p) = \text{tr}(p) = 0$ and $2\omega_2(p) = (\text{tr}(p))^2 - \text{tr}(p^2) = 0$.

Hence $\omega_1 = \omega_2 = 0$ if $\text{char}(K) \neq 2$; in this case ω_3 is either 0 (and hence p is PI) or not 0 (and hence p is 3-scalar).

So assume that $\text{char}(K) = 2$. Recall that

$$0 = \text{tr}(p^3) = \lambda_1^3 + \lambda_2^3 + \lambda_3^3 = \lambda_1^3 + \lambda_2^3 + \lambda_3^3 - 3\lambda_1\lambda_2\lambda_3 + 3\lambda_1\lambda_2\lambda_3.$$

But $\lambda_1^3 + \lambda_2^3 + \lambda_3^3 - 3\lambda_1\lambda_2\lambda_3$ is a multiple of $\lambda_1 + \lambda_2 + \lambda_3$ (seen by substituting $-(\lambda_1 + \lambda_2)$ for λ_3) and thus equals 0. Thus, $0 = 3\lambda_1\lambda_2\lambda_3 = \lambda_1\lambda_2\lambda_3 = \omega_3(p)$, and the Hamilton–Cayley equation yields $p^3 + \omega_2 p = 0$. Therefore, $p(p^2 + \omega_2) = 0$ and by Amitsur's theorem either p is PI, or $p^2 = -\omega_2$ (which is central), implying by Lemma 1.33 that p is central. ■

3.3.2 Evaluations of multilinear trace zero polynomials on 3×3 matrices

Proof of Theorem 1.19. If the polynomial $\omega_2(p)$ (defined in the proof of Theorem 1.18) is identically zero, then the characteristic polynomial is $p^3 - \omega_3(p) = 0$, implying p is either scalar (which can happen only if $\text{Char}(K) = 3$) or 3-scalar. Therefore we may assume that the polynomial $\omega_2(p)$ is not identically zero. Let

$$f_{\alpha,\beta}(M) = \alpha\omega_2(M)^3 + \beta\omega_3(M)^2.$$

It is enough to show that for any $\alpha, \beta \in K$ there exists a non-nilpotent matrix $M = p(a_1, \dots, a_m)$ such that $f_{\alpha,\beta}(p(a_1, \dots, a_m)) = 0$, since this will imply that the image of H (defined in Lemma 3.1) contains all $-\frac{\beta}{\alpha}$ and thus $K \cup \{\infty\}$. (For example, if $\alpha = 0$ and $\beta \neq 0$, then $\omega_3(M) = 0$, implying $\omega_2(M) \neq 0$ since $\omega_1(M) = 0$ and M is non-nilpotent, and thus $H = \infty$.) Therefore, for any trace vanishing polynomial (i.e., a polynomial $x^3 + \gamma_1 x + \gamma_0$) there is a matrix in $\text{Im } p$ for which this is the characteristic polynomial. Hence whenever $\lambda_1 + \lambda_2 + \lambda_3 = 0$ there is a matrix with eigenvalues λ_i .

We may assume that $a = p(Y_1, \dots, Y_m)$ and $b = p(\tilde{Y}_1, Y_2, \dots, Y_m)$ are not proportional, for generic matrices $\tilde{Y}_1, Y_1, \dots, Y_m$, cf. Lemma 1.34. Consider the polynomial $\varphi_{\alpha,\beta}(t) = f_{\alpha,\beta}(a + tb)$. There are three cases to consider:

Case I. $\varphi_{\alpha,\beta} = 0$ identically. Then $f_{\alpha,\beta}(a) = 0$, and a is not nilpotent by Proposition 1.31. Here we use the Deligne trick.

Case II. $\varphi_{\alpha,\beta}$ is a constant $\tilde{\beta} \neq 0$. Then $f_{\alpha,\beta}(b + ta) = t^6 \varphi_{\alpha,\beta}(t^{-1}) = \tilde{\beta} t^6$; thus $f_{\alpha,\beta}(b) = 0$, and b is not nilpotent by Proposition 1.31. Here we use the Deligne trick.

Case III. $\varphi_{\alpha,\beta}$ is not constant. Then it has finitely many roots. Assume that for each substitution t the matrix $a + tb$ is nilpotent; in particular, $\omega_2(a + tb) = 0$. Note that $\omega_2(a + tb)$ equals the sum of principal 2×2 minors and thus is a quadratic polynomial (for otherwise

$\omega_2(b) = 0$ which means that $\omega_2(p)$ is identically zero, a contradiction). Hence $\omega_2(a + tb)$ has two roots, which we denote as t_1 and t_2 . If $t_1 = t_2$, then t_1 is uniquely defined and thus, in view of Remark 1.42, is a rational function in the entries of a and b , and $a + t_1b$ is a nilpotent rational function (because we assumed that one of $a + t_1b$ and $a + t_2b$ is nilpotent, but here they are equal.) At least one of t_1 and t_2 is a root of $\varphi_{\alpha,\beta}$.

If only t_1 is a root, then t_1 is uniquely defined and thus, by Remark 1.42, is a rational function; hence, $a + t_1b$ is a nilpotent polynomial, contradicting Proposition 1.31. Here we use the Deligne trick. Thus, we may assume that both t_1 and t_2 are roots of $\varphi_{\alpha,\beta}$. But $\varphi_{\alpha,\beta}(t_i)$ is nilpotent, and in particular $\omega_3(a + t_ib) = 0$. Thus there exists exactly one more root t_3 of $\omega_3(a + tb)$, which is uniquely defined and thus, by Remark 1.42, is rational. Hence we may consider the polynomial $q(x_1, \dots, x_m, \tilde{x}_1) = a + t_3b$, which must satisfy the condition $\text{tr}(q) = \det(q) = 0$. This is impossible for homogeneous q by Theorem 1.18, and also impossible for nonhomogeneous q since the leading homogenous component q_d would satisfy $\text{tr}(q_d) = \det(q_d) = 0$, a contradiction. ■

3.4 The Euler graph approach

We recall the following elementary graph-theoretic Lemma 1.27: *Let p be a multilinear polynomial. If the a_i are matrix units, then $p(a_1, \dots, a_m)$ is either 0, or $c \cdot e_{ij}$ for some $i \neq j$, or a diagonal matrix.*

Proof. We connect a vertex i with a vertex j by an oriented edge if there is a matrix e_{ij} in our set $\{a_1, a_2, \dots, a_m\}$. The evaluation $p(a_1, \dots, a_m) \neq 0$ only if there exists an Eulerian cycle or an Eulerian path in the graph. There exists an Eulerian path only if the degrees of all vertices but two are even, and the degrees of these two vertices are odd. Also we know that there exists an Eulerian cycle only if the degrees of all vertices are even. Thus when $p(a_1, \dots, a_m) \neq 0$, there exists either an Eulerian path or cycle in the graph. In the first case we have exactly two vertices of odd degree such that one of them (i) has more output edges and another (j) has more input edges. Thus the only nonzero terms in the sum of our polynomial can be of the type ce_{ij} and therefore the result will also be of this type. In the second case all degrees are even. Thus there are only cycles and the result must be a diagonal matrix. ■

Theorem 3.10. *If there exist α , β , and γ in K such that $\alpha + \beta + \gamma$, $\alpha + \beta\varepsilon + \gamma\varepsilon^2$ and $\alpha + \beta\varepsilon^2 + \gamma\varepsilon$ are nonzero, together with matrix units e_1, e_2, \dots, e_m such that $p(e_1, e_2, \dots, e_m)$ has eigenvalues α , β and γ , then $\text{Im } p$ is dense in M_3 .*

Proof. Define χ to be the permutation of the set of matrix units, sending the indices $1 \rightarrow 2$, $2 \rightarrow 3$, and $3 \rightarrow 1$. For example, $\chi(e_{12}) = e_{23}$. For triples T_1, \dots, T_m (each $T_i = (t_{i,1}, t_{i,2}, t_{i,3})$) consider the function

$$\begin{aligned} f(T_1, \dots, T_m) = & p(t_{1,1}x_1 + t_{1,2}\chi(x_1) + t_{1,3}\chi^{-1}(x_1), t_{2,1}x_2 + t_{2,2}\chi(x_2) + t_{2,3}\chi^{-1}(x_2), \\ & \dots, t_{m,1}x_m + t_{m,2}\chi(x_m) + t_{m,3}\chi^{-1}(x_m)). \end{aligned} \quad (3.1)$$

Opening the brackets, we have 3^m terms, each of which we claim is a diagonal matrix. Each term is a monomial with coefficient of the type

$$\chi^{k_{\pi(1)}} \chi^{k_{\pi(2)}} \dots \chi^{k_{\pi(m)}} x_{\pi(1)} x_{\pi(2)} \dots x_{\pi(m)},$$

where k_i is -1 , 0 or 1 , and π is a permutation. Since we substitute only matrix units in p , by Lemma 1.27 the image is either diagonal or a matrix unit with some coefficient. For each of the three vertices v_1, v_2, v_3 in our graph define the index ι_ℓ , for $1 \leq \ell \leq 3$ to be the number of incoming edges to v_ℓ minus the number of outgoing edges from v_ℓ . Thus, at the outset, when the image is diagonal, we have $\iota_1 = \iota_2 = \iota_3 = 0$.

We claim that after applying χ to any matrix unit the new ι'_ℓ will all still be congruent modulo 3. Indeed, if the edge $\vec{12}$ is changed to $\vec{23}$, then $\iota'_1 = \iota + 1$ and $\iota'_3 = \iota_3 + 1$, whereas $\iota'_2 = \iota_2 - 2 \equiv \iota_2 + 1$. The same with changing $\vec{23}$ to $\vec{31}$ and $\vec{31}$ to $\vec{12}$. If we make the opposite change $\vec{21}$ to $\vec{13}$, then (modulo 3) we subtract 1 throughout. If we make a change of the type $\vec{ii} \mapsto \vec{jj}$, then $\iota'_\ell = \iota_\ell$ for each ℓ .

If $p(\chi^{k_1}x_1, \chi^{k_2}x_2, \dots, \chi^{k_m}x_m) = e_{ij}$, this means that the number of incoming edges minus the number of outgoing edges of the vertex i is $-1 \pmod{3}$ and the number of incoming edges minus the number of outgoing edges of j is $1 \pmod{3}$, which are not congruent modulo 3. Thus the values of the mapping f defined in (3.1) are diagonal matrices. Now fix $3m$ algebraically independent triples $T_1, \dots, T_m, \Theta_1, \dots, \Theta_m, \Upsilon_1, \dots, \Upsilon_m$. Assume that $\text{Im } f$ is 2-dimensional. Then $\text{Im } df$ must also be 2-dimensional at any point. Consider the differential df at the point $(\Theta_1, T_2, \dots, T_m)$. Thus,

$$f(\Theta_1, T_2, \dots, T_m), \quad f(T_1, T_2, \dots, T_m), \quad f(\Theta_1, \Theta_2, \dots, T_m)$$

belong to $\text{Im } df$. Thus these three matrices must span a linear space of dimension not more than 2. Hence they lie in some plane P . Now take

$$f(\Theta_1, \Theta_2, T_3, \dots, T_m), \quad f(\Theta_1, T_2, T_3, \dots, T_m), \quad f(\Theta_1, \Theta_2, \Theta_3, T_4, \dots, T_m).$$

For the same reason they lie in a plane, which is the plane P because it has two pure quaternions from P . By the same argument, we conclude that all the matrices of the type $f(\Theta_1, \dots, \Theta_k, T_{k+1}, \dots, T_m)$ lie in P . Now we see that

$$f(\Theta_1, \dots, \Theta_{m-1}, T_m), \quad f(\Theta_1, \dots, \Theta_m), \quad f(\Upsilon_1, \Theta_2, \dots, \Theta_m)$$

also lie in P . Analogously we obtain that also

$$f(\Upsilon_1, \dots, \Upsilon_k, \Theta_{k+1}, \dots, \Theta_m) \in P$$

for any k .

Hence for $3m$ algebraically independent triples

$$T_1, \dots, T_m; \quad \Theta_1, \dots, \Theta_m; \quad \Upsilon_1, \dots, \Upsilon_m,$$

we have obtained that $f(T_1, \dots, T_m)$, $f(\Theta_1, \dots, \Theta_m)$ and $f(\Upsilon_1, \dots, \Upsilon_m)$ lie in one plane. Thus any three values of f , in particular $\text{Diag}\{\alpha, \beta, \gamma\}$, $\text{Diag}\{\beta, \gamma, \alpha\}$ and $\text{Diag}\{\gamma, \alpha, \beta\}$, must lie in one plane. We claim that this can happen only if

$$\alpha + \beta + \gamma = 0, \quad \alpha + \beta\varepsilon + \gamma\varepsilon^2 = 0, \quad \text{or} \quad \alpha + \beta\varepsilon^2 + \gamma\varepsilon = 0.$$

Indeed, $\text{Diag}\{\alpha, \beta, \gamma\}$, $\text{Diag}\{\beta, \gamma, \alpha\}$ and $\text{Diag}\{\gamma, \alpha, \beta\}$, are dependent if and only if the matrix

$$\begin{pmatrix} \alpha & \beta & \gamma \\ \beta & \gamma & \alpha \\ \gamma & \alpha & \beta \end{pmatrix}$$

is singular, i.e., its determinant $3\alpha\beta\gamma - (\alpha^3 + \beta^3 + \gamma^3) = 0$. But this has the desired three roots when viewed as a cubic equation in γ .

We have a contradiction to our hypothesis. ■

Remark 3.11. If there exist α, β , and γ such that $\alpha + \beta + \gamma = 0$ but (α, β, γ) is not proportional to $(1, \varepsilon, \varepsilon^2)$ or $(1, \varepsilon^2, \varepsilon)$, with matrices A_1, A_2, \dots, A_m such that $p(A_1, A_2, \dots, A_m)$ has eigenvalues α, β and γ , then either all diagonalizable trace zero matrices lie in $\text{Im } p$, or $\text{Im } p$ is dense in $M_3(K)$. If $\alpha + \beta\varepsilon + \gamma\varepsilon^2 = 0$ but (α, β, γ) is not proportional to $(1, \varepsilon, \varepsilon^2)$ or $(1, 1, 1)$, then all diagonalizable matrices with eigenvalues $\alpha + \beta, \alpha + \beta\varepsilon$ and $\alpha + \beta\varepsilon^2$ lie in $\text{Im } p$ or $\text{Im } p$ is dense in $M_3(K)$.

Remark 3.12. The proof of Theorem 3.10 works also for any field K of characteristic 3. In this case $\varepsilon = 1$. Hence, if there are α, β , and γ in K such that

$$\alpha + \beta + \gamma \neq 0,$$

together with matrix units e_1, e_2, \dots, e_m such that $p(e_1, e_2, \dots, e_m)$ has eigenvalues α, β and γ , then $\text{Im } p$ is dense in M_3 . Therefore, for $\text{Char } K = 3$, any multilinear polynomial p is either trace vanishing or $\text{Im } p$ is dense in $M_3(K)$.

3.5 Open problems related to the rank 3 case

Problem 3.13. *Does there actually exist a multilinear polynomial whose image evaluated on 3×3 matrices consists of 3-scalar matrices?*

Problem 3.14. *Does there actually exist a multilinear polynomial whose image evaluated on 3×3 matrices is the set of scalars plus 3-scalar matrices?*

Remark 3.15. Problems 3.13 and 3.14 both have the same answer. If they both have affirmative answers, such a polynomial would provide a counter-example to Kaplansky's problem.

Problem 3.16. *Is it possible that the image of a multilinear polynomial evaluated on 3×3 matrices is dense but not all of $M_3(K)$?*

Problem 3.17. *Is it possible that the image of a multilinear polynomial evaluated on 3×3 matrices is the set of all trace vanishing matrices excluding discriminant vanishing diagonalizable matrices?*

Problem 3.18. *Give a classification of all possible evaluations of homogeneous polynomials on $M_3(K)$ with respect to Zariski closure.*

Remark 3.19. The working hypothesis is that there are 6 Zariski closures of image sets of homogeneous polynomials: $\{0\}, K, V_3(K), \text{sl}_3(K), V_3+K, M_3(K)$. However the problem remains open.

Problem 3.20. *Investigate all possible image sets of non homogeneous polynomials in order to obtain a zoo of interesting examples.*

Example 3.21. Here are some interesting examples.

- Let $p(x_1, \dots, x_m)$ be a homogeneous 3-central polynomial and $h(x)$ be any polynomial in one variable. Then, $h(p(x_1, \dots, x_m))$ is the polynomial having evaluations with triples of eigenvalues belonging to the set $(h(c), h(c\varepsilon), h(c\varepsilon^2))$ for all $c \in K$. This is a 7-dimensional image.
- Let $h(x)$ be again an arbitrary polynomial in one variable, and consider the polynomial $h([x, y])$. This polynomial image is the set of all matrices having triples of eigenvalues belonging to the set $(h(a), h(b), h(-a-b))$ for all pairs $a, b \in K$. This is an 8-dimensional image.

Problem 3.22. *Give a classification of all possible evaluations of multilinear polynomials on $M_3(\mathbb{R})$ with respect to the standard topology.*

Remark 3.23. Although we cannot answer a question about existence of multilinear 3-central polynomials, but we know that multilinear 3-central polynomials with real coefficients do not exist. Indeed, according to Lemma 3.3 any multilinear polynomial has a diagonal not scalar matrix evaluation (with real entries), and these real entries are its eigenvalues. Note that a nonzero 3-central matrix cannot have three real eigenvalues.

Problem 3.24. *Investigate possible evaluations of multilinear polynomials on the simple non special Jordan algebra, which is denoted as $H(C_3)$ in [109] and C_{27} in [100].*

4 The case of high rank

The Euler graph approach is useful also for $n > 3$:

4.1 The Euler graph approach for algebras of high rank

4.1.1 The idea of the method

Define χ to be the permutation of the set of matrix units, sending the index $i \mapsto i + 1$ for $1 \leq i \leq n - 1$, and $n \mapsto 1$. For example, $\chi(e_{12}) = e_{23}$, $\chi(e_{57}) = e_{68}$. Since we substitute only matrix units into p , Lemma 1.27 shows that the image is either diagonal or a matrix unit with some coefficient. Consider the corresponding graph Γ . Consider the sum of $(i - j)$ over edges $(i \rightarrow j)$ of Γ . If the graph is an Eulerian cycle then this sum is 0, and if it is an Eulerian path from k to ℓ then this sum equals $k - \ell$. Take matrix units a_1, \dots, a_m such that $p(a_1, \dots, a_m)$ is a diagonal matrix for some $\alpha \in K$. We may assume that $a_1 \cdots a_m = D$. Writing $a_\ell = e_{i_\ell, j_\ell}$, we define $\iota(a_\ell) = i_\ell - j_\ell$. Thus $i_1 = j_m$ and $\sum \iota(a_\ell) = 0$. Then $\chi^k(a_\ell) = e_{i_\ell+k, j_\ell+k}$ (taken modulo n), implying

$$\iota(\chi^k(a_\ell)) \equiv (i_\ell + k) - (j_\ell + k) = i_\ell - j_\ell = \iota(a_\ell) \pmod{n}.$$

Consider

$$f(a_1, \dots, a_m) = p \left(\sum_{k_1=1}^m t_{k_1,1} \chi^{k_1}(a_1), \dots, \sum_{k_m=1}^m t_{k_m,m} \chi^{k_m}(a_m) \right), \quad (4.1)$$

where the $t_{k,\ell}$ are commuting indeterminates. Opening the brackets, we have n^m terms, each of the form

$$a' = \chi^{k_1}(a_{\pi(1)}) \cdots \chi^{k_m}(a_{\pi(m)}),$$

which, if nonzero, must have

$$\iota(a') \equiv \sum_{\ell=1}^m \iota(\chi^{k_\ell} a_{\pi(\ell)}) \equiv \sum_{\ell=1}^m \iota(a_{\pi(\ell)}) \equiv 0 \pmod{n},$$

implying a' is a diagonal matrix. Hence $\text{Im } f \subseteq \text{Im } p$ contains only diagonal matrices. This helps us to obtain a good subset in the set of diagonal properties, in particular matrices with needed set of eigenvalues, and each obtained matrix is an evaluation of the polynomial p .

4.1.2 Applications of the method

In this subsection we will show how this method is used in the proof of several important results: Theorems 1.12, 1.21 and 1.22.

Proof of Theorem 1.12. Define χ as in Section 4.1.1 By Lemma 1.27 there exist matrix units a_1, \dots, a_m with $p(a_1, \dots, a_m) = \alpha e_{12}$ for some $\alpha \in K$. We may assume that $a_1 \cdots a_m = e_{12}$. Writing $a_\ell = e_{i_\ell, j_\ell}$, we define $\iota(a_\ell) = i_\ell - j_\ell$. Thus $i_1 = 1$ and $j_m = 2$, and $\sum \iota(a_\ell) = 1$. Then $\chi^k(a_\ell) = e_{i_\ell+k, j_\ell+k}$ (taken modulo n), implying

$$\iota(\chi^k(a_\ell)) \equiv (i_\ell + k) - (j_\ell + k) = i_\ell - j_\ell = \iota(a_\ell) \pmod{n}.$$

Consider $f(a_1, \dots, a_m)$ defined in (4.1). Opening the brackets, we have n^m terms, each of the form

$$a' = \chi^{k_1}(a_{\pi(1)}) \cdots \chi^{k_m}(a_{\pi(m)}),$$

which, if nonzero, must have

$$\iota(a') \equiv \sum_{\ell=1}^m \iota(\chi^{k\ell} a_{\pi(\ell)}) \equiv \sum_{\ell=1}^m \iota(a_{\pi(\ell)}) \equiv 1 \pmod{n},$$

implying a' is a matrix of the form $ce_{i,i+1}$ or $ce_{n,1}$. Hence $\text{Im } f \subseteq \text{Im } p$ has the form

$$a = \begin{pmatrix} 0 & * & 0 & \dots & 0 \\ 0 & 0 & * & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & * \\ * & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Each of the starred entries of a is a polynomial with respect to $t_{k,i}$ and each of them takes nonzero values because $e_{k,k+1}$ belongs to the image of f for any $1 \leq k \leq n-1$ and also $e_{n,1} \in \text{Im}(f)$. Therefore for generic $t_{k,\ell}$ each of the starred entries is nonzero, so the minimal polynomial of a is $\lambda^n - \alpha$ for some α , implying a has eigenvalues $\{c, c\varepsilon, \dots, c\varepsilon^{n-1}\}$ where c is the n -th root of the determinant α . \blacksquare

Remark 4.1. Assume that K has the form $F[\varepsilon]$, where ε is a primitive n -th root of 1. Let $\bar{}$ denote the automorphism of K sending $\varepsilon \mapsto \varepsilon^{-1}$.

Let us introduce the tool of “harmonic bases” of the space of diagonal matrices. There is a base of the matrices e_k for $0 \leq k \leq n-1$, where $e_k = \text{Diag}\{1, \varepsilon^k, \varepsilon^{2k}, \dots, \varepsilon^{(n-1)k}\}$. Assume that there exist matrix units a_1, \dots, a_m such that $p(a_1, \dots, a_m) = \text{Diag}\{c_0, \dots, c_{n-1}\}$. This can be written as a linear combination of the e_k . If, for some k , e_k appears in this sum with nonzero coefficient, then e_k belongs to the linear span of $\langle M, \chi(M), \chi^2(M), \dots \rangle$, where $\chi(\text{Diag}\{\lambda_1, \dots, \lambda_n\}) = \text{Diag}\{\lambda_n, \lambda_1, \dots, \lambda_{n-1}\}$. Therefore if we have a set of matrix units a_i such that $p(a_1, \dots, a_m) = M$, then we construct a multilinear mapping f whose image will either have at least dimension 3 or the image will be a linear set and therefore M will be a linear combination of no more than two base elements e_k .

Assume also that the image of p is at most $(n^2 - n + 2)$ -dimensional. By Remark 1.51, the image of f constructed in the proof of Theorem 1.12 is at most 2-dimensional and thus is a linear space. If $p(a_1, \dots, a_m) = h_0 e_0 + h_1 e_1 + \dots + h_{n-1} e_{n-1}$ with $h_k \neq 0$, then e_k belongs to the linear span of $\text{Im } f$. Hence there are at most two nonzero coefficients, say, h_k and h_l with all of the others zero. We can consider the scalar product

$$\langle \{\alpha_1, \dots, \alpha_n\} \{\beta_1, \dots, \beta_n\} \rangle = \sum_{i=1}^n \alpha_i \bar{\beta}_i.$$

We compute $\langle \{c_0, \dots, c_{n-1}\}, e_s \rangle$ in two ways, first as nq_s and then as

$$c_0 + c_1 \varepsilon^{-s} + \dots + c_{n-1} \varepsilon^{-(n-1)s}$$

since $\bar{\varepsilon^l} = \varepsilon^{-l}$.

Proof Theorem 1.21. There exist matrix units a_i such that $p(a_1, \dots, a_m) = \text{Diag}\{c_0, \dots, c_{n-1}\}$ is diagonal but not scalar. If at least one of the c_i were zero, then each c_i would be zero and therefore this matrix is zero, a contradiction. Thus, without loss of generality $c_0 \neq c_1$ are nonzero. Therefore there also exist matrix units \tilde{a}_i such that $p(\tilde{a}_1, \dots, \tilde{a}_m) = \text{Diag}\{c_1, c_0, c_2, \dots, c_{n-1}\}$.

Assume that the image of f as constructed in (4.1) is 1-dimensional. Then, for each i , $\text{Diag}\{c_0, \dots, c_{n-1}\}$ is proportional to $\text{Diag}\{c_i, c_{i+1}, \dots, c_{n-1}, c_0, \dots, c_{i-1}\}$. We can construct the mappings f and \tilde{f} as before, and their images cannot be both 1-dimensional since otherwise

$$\tau = \frac{c_1}{c_0} = \frac{c_2}{c_1} = \frac{c_3}{c_2} = \dots = \frac{c_n}{c_{n-1}} = \frac{c_0}{c_n},$$

and also $\tilde{\tau} = \frac{c_0}{c_1} = \frac{c_2}{c_0} = \frac{c_3}{c_2} = \dots$. Hence, since $n - 1 \geq 3$.

$$\tau^2 = \frac{c_2}{c_1} \cdot \frac{c_1}{c_0} = \frac{c_2}{c_0} = \tilde{\tau} = \frac{c_3}{c_2} = \tau.$$

Thus $\tau \in \{0, 1\}$. If $\tau = 1$ then $p(a_1, \dots, a_m)$ is scalar, a contradiction. If $\tau = 0$ then $c_1 = 0$, a contradiction.

We conclude that $\text{Im } f$ is least 2-dimensional and $\text{Im } p$ is at least $(n^2 - n + 2)$ -dimensional.

Assume that the image of p is at most $(n^2 - n + 2)$ -dimensional. As we showed in Remark 4.1, the matrix $\text{Diag}\{c_0, \dots, c_{n-1}\}$ can be written as $\alpha e_k + \beta e_l$, which is not scalar. Without loss of generality we may assume that $c_0 \neq c_1$ (because there exists r such that $c_r \neq c_{r+1}$), and we now consider the matrix

$$\text{Diag}\{c_r, c_{r+1}, \dots, c_{n-1} c_0, c_1, \dots, c_{r-1}\}$$

with its different coefficients $\tilde{\alpha} = \varepsilon^{rk} \alpha$ and $\tilde{\beta} = \varepsilon^{rl} \beta$.

We define the matrices $q_k := \text{Diag}\{\varepsilon^k, 1, \varepsilon^{2k}, \varepsilon^{3k}, \dots, \varepsilon^{(n-1)k}\}$. Switching the indices 1 and 2, we obtain matrix units \tilde{a}_i such that

$$p(\tilde{a}_1, \dots, \tilde{a}_m) = \text{Diag}\{c_1, c_0, c_2, c_3, \dots, c_{n-1}\} = \alpha q_k + \beta q_l.$$

By Remark 4.1, $\alpha q_k + \beta q_l$ also can be written as a linear combination of two elements of the base e_s (say, $\tilde{\alpha} e_{\tilde{k}} + \tilde{\beta} e_{\tilde{l}}$). Note that

$$\langle q_k, e_s \rangle = \varepsilon^k + \varepsilon^{-s} - 1 - \varepsilon^{k-s} = (\varepsilon^k - 1)(1 - \varepsilon^{-s}) + \langle e_k, e_s \rangle.$$

Thus, if $k \neq s$, then

$$\langle q_k, e_s \rangle = \varepsilon^k + \varepsilon^{-s} - 1 - \varepsilon^{k-s} = (\varepsilon^k - 1)(1 - \varepsilon^{-s})$$

since $\langle e_k, e_s \rangle = 0$, and if $k = s$ then

$$\langle q_k, e_s \rangle = (\varepsilon^k - 1)(1 - \varepsilon^{-s}) + n.$$

Hence, if $s \notin \{k, l\}$ then

$$\langle \alpha q_k + \beta q_l, e_s \rangle = (1 - \varepsilon^{-s})(\alpha(\varepsilon^k - 1) + \beta(\varepsilon^l - 1)).$$

We denote $\delta = \alpha(\varepsilon^k - 1) + \beta(\varepsilon^l - 1)$. Recall that $c_1 \neq c_0$, and thus $\delta \neq 0$. Therefore either $s = \tilde{k}$, or else $s = \tilde{l}$, or $\langle \alpha q_k + \beta q_l, e_s \rangle = 0$ (and thus s is either k , or l , or $1 - \varepsilon^{-s} = 0$ (and thus $s = 0$) – only five possibilities. But for $n \geq 6$ there are at least three nonzero coefficients, a contradiction.

Thus we may assume that $n = 5$. We have exactly five possibilities for s , which therefore must be distinct. Therefore the k -th and l -th coefficients of $\alpha q_k + \beta q_l$ will be zero, i.e.,

$$(1 - \varepsilon^{-k})\delta + 5\alpha = (1 - \varepsilon^{-l})\delta + 5\beta = 0,$$

where $\delta = \alpha(\varepsilon^k - 1) + \beta(\varepsilon^l - 1)$. In particular

$$\frac{\alpha}{\beta} = \frac{1 - \varepsilon^{-k}}{1 - \varepsilon^{-l}}.$$

Now let us take matrix units a'_i such that $p(a'_1, \dots, a'_m) = \text{Diag}\{c_2, c_1, c_0, c_3, c_4\}$. Then $\alpha r_k + \beta r_l$ can also be written as a linear combination of two of the e_s , where $r_k = \text{Diag}\{\varepsilon^{2k}, \varepsilon^k, 1, \varepsilon^{3k}, \varepsilon^{4k}\}$. If $k \neq s$, then

$$\langle r_k, e_s \rangle = \varepsilon^{2k} + \varepsilon^{-2s} - 1 - \varepsilon^{2k-2s} = (\varepsilon^{2k} - 1)(1 - \varepsilon^{-2s}).$$

We perform the same calculations as before, and obtain

$$\frac{\alpha}{\beta} = \frac{1 - \varepsilon^{-2k}}{1 - \varepsilon^{-2l}}.$$

Hence

$$\frac{1 - \varepsilon^{-k}}{1 - \varepsilon^{-l}} = \frac{1 - \varepsilon^{-2k}}{1 - \varepsilon^{-2l}},$$

implying

$$\frac{1 + \varepsilon^{-k}}{1 + \varepsilon^{-l}} = 1,$$

and hence $k = l$, a contradiction. ■

Proof of Theorem 1.22. First note that $4^2 - 4 + 2 = 14$, so $\dim \text{Im } p \geq 14$. Assume that p is a multilinear polynomial evaluated on 4×4 matrices with 14-dimensional image. Let a_1, \dots, a_m be any matrix units such that $p(a_1, \dots, a_m)$ is diagonal but not scalar. Let $p(a_1, \dots, a_m) = \text{Diag}\{c_0, c_1, c_2, c_3\}$ and $c_0 \neq c_1$. We use the same notation as in the proof of Theorem 1.22. Recall that $e_k = \text{Diag}\{1, i^k, i^{2k}, i^{3k}\}$ and $q_k = \text{Diag}\{i^k, 1, i^{2k}, i^{3k}\}$. As in the proof of Theorem 1.22, $\langle \alpha q_k + \beta q_l, e_s \rangle = \delta(1 - i^{-s})$ if $s \notin \{k, l\}$, or $\delta(1 - i^{-s}) + 4\alpha$ if $s = k$ and $\delta(1 - i^{-s}) + 4\beta$ if $s = l$. Therefore k and l are nonzero (for otherwise we have two nonzero possibilities for $s \notin \{k, l\}$ and one other nonzero coefficient would be zero: $\langle \alpha q_k + \beta q_l, e_0 \rangle = 4\alpha$ (if we assume $k = 0$ without loss of generality)). Therefore $p(a_1, \dots, a_m)$ belongs to the linear span $\langle e_1, e_2, e_3 \rangle$. Hence we have three options:

- $p(a_1, \dots, a_m) = \alpha e_1 + \beta e_2$,
- $p(a_1, \dots, a_m) = \alpha e_1 + \beta e_3$,
- $p(a_1, \dots, a_m) = \alpha e_3 + \beta e_2$.

We will not treat the last case since its calculations are as in the first case. Let us consider the first case $p(a_1, \dots, a_m) = \alpha e_1 + \beta e_2$. Therefore $p(\tilde{a}_1, \dots, \tilde{a}_m) = \alpha q_1 + \beta q_2$ which can be written explicitly as

$$\left(\frac{1}{2}\alpha - \frac{1+i}{2}\beta\right) e_1 + \frac{i-1}{2}\alpha e_2 + \left(\frac{i}{2}\alpha + \frac{i-1}{2}\beta\right) e_3,$$

thus $\alpha \in \{0, (1+i)\beta, -(1+i)\beta\}$. If $\alpha = (1+i)\beta$ then $p(a_1, \dots, a_m) = \beta \text{Diag}\{i-2, i+2, -i, -i\}$. If $\alpha = -(1+i)\beta$ then $p(a_1, \dots, a_m) = \beta \text{Diag}\{-i, -i, i+2, i-2\}$. In both cases $\alpha = \pm(1+i)\beta$, so there are matrix units \tilde{a}_i such that $p(\tilde{a}_1, \dots, \tilde{a}_m) = \text{Diag}\{-i, i+2, -i, i-2\}$ which can be written explicitly as $-ie_1 - ie_2 + ie_3$, and we have three nonzero coefficients. We conclude that

the image is at least 15-dimensional. If $\alpha = 0$, then the value $p(a_1, \dots, a_m)$ has eigenvalues $(c, c, -c, -c)$ as in the conditions of the theorem.

Assume now $p(a_1, \dots, a_m) = \alpha e_1 + \beta e_3$. Therefore $p(a_1, \dots, a_m) = \text{Diag}\{x, y, -x, -y\}$. Then we consider $p(\tilde{a}_1, \dots, \tilde{a}_m) = \text{Diag}\{x, -x, y, -y\}$ which can be written explicitly as

$$\frac{(x-y)(1+i)}{4}e_1 + \frac{x+y}{2}e_2 + \frac{(x-y)(1-i)}{4}e_3,$$

therefore $x = \pm y$ and once again we have a matrix from the conditions of the theorem.

Therefore there is a set of matrix units a_i with $p(a_1, \dots, a_m) = \text{Diag}\{c, c, -c, -c\}$. Now we construct the mapping f and the image will be 2-dimensional if and only if it is the set $\text{Diag}\{\lambda_1, \lambda_2, -\lambda_1, -\lambda_2\}$. Therefore $\text{Im } p$ contains all the matrices with such eigenvalues, which is a 14-dimensional variety. Hence, if $\dim \text{Im } p = 14$, then $\text{Im } p$ is exactly this variety. ■

4.1.3 Questions related to the Euler graph approach

By Lemma 1.34, the image of any multilinear polynomial either is a vector space or is at least 3-dimensional. This is why, when in Section 3.4 we considered a 3-dimensional case and we obtained the needed result up to Zariski closure. However the question of using this approach for dimension larger than 3 remains open:

Question 4.2. *Is it possible to use this method in order to enlarge the dimension investigated?*

Assume we have a periodic sequence with period n (i.e., sequence $\{a_k\}$ such that $a_{k+n} = a_k$ for all $k \in \mathbb{N}$). The dimension of the space spanned by its shifts (i.e., sequences $\{b_k = a_{k+i}\}$, $i = n$ equals the number of harmonics in its discrete Fourier transform. From the other side, using conjugation we can permute numbers in the period, obtaining different periodic sequences.

Question 4.3. *What is the maximal possible number of harmonics one can obtain for a fixed sequence?*

4.2 Non-trivial images: power-central polynomials on matrices

Images of noncommutative polynomials can be nontrivial and investigation of them requires advanced theory. As a good example of such polynomials one can consider power-central polynomials.

Note that a problem of existence of 3-central polynomials occurred in Section 3.

Any multilinear non-central polynomial p (in several noncommuting variables) takes on values of degree n in the matrix algebra $M_n(F)$ over an infinite field F is called power central.

These considerations motivate the following application:

Definition 4.4. A polynomial p is ν -central if p^ν is central, for $\nu \geq 1$ minimal such. The polynomial p is *power-central* if p is ν -central, for some $\nu > 1$.

Remark 4.5. The existence of a ν -central polynomial is equivalent to UD containing an element whose ν -power is central, with ν minimal such. On the other hand, any such element can be specialized to an arbitrary division algebra of dimension n^2 over its center. Thus, to prove the non-existence of a ν -central polynomial, it suffices for suitable ℓ dividing ν to construct a division algebra with center $\hat{K} \supseteq K$ having the property that if $d^\nu \in \hat{K}$ then $d^\ell \in \hat{K}$.

Examining dimensions of images, we conclude from Theorem 1.21 that multilinear power-central polynomials do not exist whenever $n \geq 4$. Note that it does not contradict to the Theorem 4.7, where we will show that if a multilinear polynomial p is ν -central on $M_n(K)$ then $\nu = n$. Both statements hold. What we can conclude that power central multilinear polynomials

evaluated on matrix algebras $M_n(K)$ can occur only for $n = 2$ or $n = 3$. For $n = 2$ there exists an example (in particular, the polynomial $p(x, y) = [x, y]$, see Example 1.6(ii) for details), for $n = 3$ the question whether such polynomial exists remains being open (see Problem 3.13).

To put these results into perspective, we also consider the existence of ν -central polynomials which need not be multilinear.

Using the structure theory of division algebras, Saltman [95] proved that in characteristic 0, ν -central polynomials do not exist for odd $\nu > 1$ unless n is prime.

For $n = \nu$ prime, as explained in [93, Theorem 3.2.20], this is equivalent to Amitsur's generic division algebra being cyclic, one of the major open questions in the theory of division algebras. Although we have nothing more to say about this question, we have results for other cases.

Remark 4.6. A homogeneous 3-central polynomial for $n = 3$ was constructed in [93, Theorem 3.2.21], and a homogeneous 2-central polynomial for $n = 4$ was constructed in [93, Proposition 3.2.24].

Theorem 4.7. *If a multilinear polynomial p is ν -central on $M_n(K)$ then $\nu = n$.*

Proof. By Theorem 1.12, $\text{Im } p$ contains a matrix M of type $c_n e_{n,1} + \sum_{i=1}^{n-1} c_i e_{i,i+1}$ where $c_1 \cdots c_n \neq 0$. Hence $n \leq \nu$. To show that $n = \nu$, consider the continuous mapping

$$\theta: (x_1, \dots, x_m) \mapsto (\lambda_1 : \lambda_2 : \cdots : \lambda_n)$$

to the projective variety, where λ_i are the eigenvalues of $p(x_1, \dots, x_m)$. Since p is ν -central, θ is discrete and therefore constant. By Theorem 1.12 there are matrices a_1, \dots, a_m such that $p(a_1, \dots, a_m) = M$ and thus $\theta(a_1, \dots, a_m) = (1 : \varepsilon : \varepsilon^2 : \cdots : \varepsilon^{n-1})$. Therefore p is n -central, as desired. ■

As a special case of Remark 1.50, if p is power-central then $\text{Im } p$ has dimension $n^2 - n + 1$.

4.2.1 Considerations arising from division algebras

A major tool is Amitsur's theorem (Proposition 1.31). Here is an easy consequence of the theory of division algebras.

Lemma 4.8. *If a polynomial p is ν -central for $M_n(K)$ for $\nu > 1$, then ν cannot be relatively prime to n .*

Proof. We can view p as an element of the generic division algebra $\widetilde{\text{UD}}$ of degree n , and we adjoin an ν -root of 1 to K if necessary. Then p generates a subfield of $\widetilde{\text{UD}}$, of dimension dividing ν . Hence the dimension is 1; i.e., $\nu = 1$. ■

We shall need a general fact about polynomial evaluations.

Lemma 4.9. *For any polynomial $p(x_1, \dots, x_m)$ which has an evaluation of degree n on $M_n(K)$, there is an index i , $1 \leq i \leq m$, and matrices $a_1, a_2, \dots, a_m, a'_i$ such that the evaluations $p(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_m)$ and $p(a_1, \dots, a_{i-1}, a'_i, a_{i+1}, \dots, a_m)$ do not commute.*

Proof. We go back and forth to generic matrices and $\widetilde{\text{UD}}$. First of all, for all generic matrices $Y_1, \dots, Y_m, Y'_1, \dots, Y'_m$, and each i , clearly $p(Y_1, \dots, Y_i, Y'_{i+1}, \dots, Y'_m)$ has degree n over K_1 , and thus has distinct eigenvalues, from which it follows at once that $p(Y_1, \dots, Y_m)$ and $p(Y'_1, \dots, Y'_m)$ do not commute (since one could diagonalize $p(Y_1, \dots, Y_m)$ while $p(Y'_1, \dots, Y'_m)$ remains non-diagonal).

But, for each i , $K_1(p(Y_1, \dots, Y_i, Y'_{i+1}, \dots, Y'_m))$ has dimension n over K_1 and thus is a maximal subfield of $\widetilde{\text{UD}}$. It follows that $p(Y_1, \dots, Y_i, Y'_{i+1}, \dots, Y'_m)$ and $p(Y_1, \dots, Y_{j-1}, Y'_j, \dots, Y'_m)$ commute iff

$$K_1(p(Y_1, \dots, Y_i, Y'_{i+1}, \dots, Y'_m)) = K_1(p(Y_1, \dots, Y_{j-1}, Y'_j, \dots, Y'_m)).$$

In particular, $K_1(p(Y_1, \dots, Y_m)) \neq K_1(p(Y'_1, \dots, Y'_m))$, implying

$$K_1(p(Y_1, \dots, Y_i, Y'_{i+1}, \dots, Y'_m)) \neq K_1(p(Y_1, \dots, Y_{i-1}, Y'_i, \dots, Y'_m))$$

for some i , and thus $p(Y_1, \dots, Y_i, Y'_{i+1}, \dots, Y'_m)$ and $p(Y_1, \dots, Y_{i-1}, Y'_i, \dots, Y'_m)$ do not commute. In other words,

$$[p(Y_1, \dots, Y_i, Y'_{i+1}, \dots, Y'_m), p(Y_1, \dots, Y_{i-1}, Y'_i, \dots, Y'_m)] \neq 0,$$

implying there is a specialization

$$Y_1 \mapsto a_1, \dots, Y_i \mapsto a_i, Y'_i \mapsto a'_i, Y'_{i+1} \mapsto a_{i+1}, \dots, Y'_m \mapsto a_m$$

yielding $[p(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_m), p(a_1, \dots, a_{i-1}, a'_i, a_{i+1}, \dots, a_m)] \neq 0$. ■

Example 4.10. Suppose K has characteristic $\neq 2$, and $n = 2^{t-1}q$ where q is odd, and construct D to be a tensor product of “generic” symbols $(\lambda_u^{n_u}, \mu_u^{n_u})$ as in [93, Example 7.1.28 and Theorem 7.1.29], where $n_1 = n_2 = \dots = n_{t-1} = 2$ and $n_t = q$. In other words, D is the algebra of central fractions of the skew polynomial ring $R := K(\varepsilon)[\lambda_u \mu_u : 1 \leq u \leq t]$ where ε is a primitive q root of 1 and the indeterminates commute except for $\lambda_u \mu_u = -\mu_u \lambda_u$ for $1 \leq u \leq t-1$ and $\lambda_t \mu_t = \varepsilon \mu_t \lambda_t$. We write a typical element of R as $\sum \alpha_{i,j} \lambda^i \mu^j$, where λ^i denotes $\prod_{u=1}^t \lambda_u^{i_u}$. There is a natural grade given by the lexicographic order on the exponents of the monomials, and it is easy to see that if $d^\nu \in K$ then the leading term $\hat{d}^\nu \in K$.

In particular, for $\nu = 2$, if $d^2 \in K$ then \hat{d} must have the form $\alpha \lambda^i \mu^j$ where $i_t = j_t = 0$. On the other hand, we claim that if $d^2 \in K$ and $\hat{d} \in K$, then $d \in K$. Indeed, taking d' to be the next leading term in d , we have

$$d^2 = (\hat{d} + d')^2 = \hat{d}^2 + 2\hat{d}d' + \dots,$$

implying $d' \in K$, and continuing, we conclude $d \in K$, as desired.

It follows that if d is 2-central then \hat{d} is 2-central.

Now we claim that D does not have 4-central elements. Indeed, if d is 4-central then d^2 is 2-central, implying \hat{d}^2 is 2-central, and thus \hat{d} is 4-central, implying \hat{d} must have the form $\alpha \lambda^i \mu^j$ where $i_t = j_t = 0$; we conclude that \hat{d} is 2-central, implying $\hat{d}^2 \in K$, and thus $d^2 \in K$ by the claim.

Theorem 4.11 (Rowen–Saltman). *4-Central polynomials do not exist.*

Proof. Combine Remark 4.5 with Example 4.10. ■

This leaves us to look for 2-central polynomials.

Proposition 4.12. *There exist homogeneous 2-central polynomials with respect to $M_n(K)$ if $n = 2q$ or $n = 4q$ for q odd.*

Proof. $\widetilde{\text{UD}}$ is a tensor product of a division algebra D_1 of degree 2 or 4 and a division algebra of degree q , and we observed earlier that D_1 has a 2-central element. ■

The situation for $8|n$ remains open, and is equivalent to another important question in division algebras about the existence of square-central elements. The following observation might be relevant, although we do not use it further.

4.2.2 2-central polynomials linear in the first indeterminate

Having settled the issue for homogeneous polynomials except for $n = 8q$, we turn to 1-linear polynomials, where the story ends differently. We use the division algebra approach.

Lemma 4.13. *Any division algebra D with a 2-central subspace V of dimension 2 contains a K -central quaternion subalgebra. In particular, $n := \deg(D)$ cannot be odd. Also, 4 does not divide n if D also has exponent n .*

Proof. Take $v, v' \in V$. Then $v^2, v'^2 \in K$. But also, by assumption, $v + v'$ is also square-central, so

$$v^2 + vv' + v'v + v'^2 = (v + v')^2 \in K,$$

implying $v'v = -vv' + \alpha$ for some $\alpha \in K$. But then $K + Kv + Kv' + Kvv'$ is a central K -subalgebra of D and has dimension at least 3, but has elements of degree 2, so has dimension 4.

The last assertion follows easily from the theory of finite dimensional division algebras. If D has a quaternion division algebra then 2 must divide n and the exponent of D is the least common multiple of 2 and $\frac{n}{2}$. ■

We are ready to improve Theorem 1.21(i) in certain cases.

Proposition 4.14. *If $p(x_1, \dots, x_m)$ is a 2-central polynomial for $n \times n$ matrices, linear in x_1 , and there are non-commuting values $p(\widetilde{a_1}, \dots, a_m)$ and $p(a'_1, \dots, a_m)$ for matrices $a_1, a'_1, a_2, \dots, a_m$, then the generic division algebra $\widetilde{\text{UD}}$ of degree n has a K_1 -central quaternion subalgebra. In particular, n cannot be odd, and 4 does not divide n .*

Proof. Let

$$w = p(Y_1, \dots, Y_m), \quad w' = p(Y'_1, Y_2, \dots, Y_m),$$

where the Y_i and Y'_1 are generic matrices. Then $w^2, w'^2 \in K_1$. But also, by definition, $w + w' = p(Y_1 + Y'_1, \dots, Y_m)$ is also 2-central, so

$$w^2 + ww' + w'w + w'^2 = (w + w')^2 \in K_1,$$

implying $w'w = -\widetilde{ww'} + \alpha$ for some $\alpha \in K_1$. But then $K_1 + K_1w + K_1w' + K_1ww'$ is a central K_1 -subalgebra of $\widetilde{\text{UD}}$ and has dimension 4.

The last assertion follows since the generic division algebra $\widetilde{\text{UD}}$ of degree n has exponent n , whereas if $\widetilde{\text{UD}}$ has a central quaternion division subalgebra, then 2 must divide n and the exponent of $\widetilde{\text{UD}}$ is the least common multiple of 2 and $\frac{n}{2}$. ■

This conclusion is the opposite of Proposition 4.12, when $n = 4q$ for q odd. The hypothesis clearly holds when p is multilinear. Indeed, in view of Lemma 4.9, two consecutive terms of the chain

$$p(a_1, \dots, a_m), p(a'_1, a_2, \dots, a_m), \dots, p(a'_1, \dots, a'_m)$$

do not commute. On the other hand, it also holds when $p(x_1, x_2, \dots, x_2)$ is non-central, since we could take $Y'_1 = Y_1 Y_2 Y_1^{-1}$.

4.3 Open problems

The investigations of possible image sets may be difficult, this is the reason to try first to construct interesting examples of polynomials with non-standard image sets:

Question 4.15. *What interesting examples of evaluations of homogeneous polynomials (with respect to Zariski topology) on high rank matrix algebras can occur?*

Remark 4.16. These examples can be nontrivial (see Section 4.2) and there is a little hope to complete description at this moment. Thus, at this moment it would be nice to provide interesting examples and constructions.

Another famous problem is the Freiheitsatz for associative algebras in nonzero characteristic:

Freiheitsatz. *For $P \in A = k\langle x_1, \dots, x_n, t \rangle$ which involves t nontrivially, the algebra $k\langle x_1, \dots, x_n \rangle$ can be naturally embedded into the quotient $k\langle x_1, \dots, x_n, t \rangle / \text{Id}(P)$.*

For $\text{Char } K = 0$ the Freiheitsatz was solved by Makar-Limanov [79] and generalized by Kolesnikov [66, 67].

Working hypothesis: The Freiheitsatz is true in arbitrary characteristic.

Methods: The Freiheitsatz for associative algebras was established in the case $\text{char}(k) = 0$ by Makar-Limanov, as a consequence of his construction of an algebraically closed skew-field. Later on, P. Kolesnikov developed Makar-Limanov's ideas and, in particular, improved his exposition. The proof is based on solvability of equations in the algebra of Malcev–von Neumann series which are related to differential operators. The construction of the algebra in which the corresponding equations were solved is none other than the $*$ -operation related to Kontsevich's formal quantization.

Makar-Limanov has proposed proving that the co-rank of an arbitrary polynomial p on $M_n(k)$ is bounded by some reasonable function of n . This would yield the Freiheitsatz almost immediately. The difficulty is that p need not be homogeneous. We have some information obtained by examining generic matrices, and the hope is that they will be amenable to new geometric techniques. Topological maps defined on Banach spaces were investigated by Brešar in [17] and his approach can be useful in this question.

4.3.1 Images of polynomial maps and matrix equations

Let P be a noncommutative polynomial.

Suppose that for all sufficiently large n we can solve the system

$$\{P_i(x_1, \dots, x_s, t) = 0\}. \quad (4.2)$$

Let $\xi_i, i = 1, \dots, s, \nu$ be a solution of (4.2). Then after substituting $x_i \mapsto a_i + \xi_i, t \mapsto t + \nu$, we obtain an equation without a free term. Further, assume that after such a substitution there appears a term $B(t)$ linear in t . Let us regard B as an element of an operator algebra. If it turns out that the operator B is invertible, we will be able to solve (4.2) using the method of consecutive approximations in the product of the matrix algebra and the free algebra, and thus see that this equation does not impose any restrictions on the original x_1, \dots, x_s .

It is worth noting that it is enough to know how to solve equations (and prove invertibility of operators) modulo matrices of bounded rank and as $n \rightarrow \infty$.

Therefore the following question is relevant:

Problem 4.17. *Give a classification of possible image sets of arbitrary (or homogeneous) polynomial evaluated on matrices of rank much higher than the polynomial degree.*

For the multilinear case, one can consider the union of all ranks matrices $M_\infty(K)$: the set of infinite matrices with finite number of nonzero entries. The working hypothesis of the L'vov–Kaplansky conjecture) is quite reasonable:

Problem 4.18. *Let p be any multilinear polynomial evaluated on $M_\infty(K)$. Then $\text{Im } p$ is the set of trace vanishing matrices $\text{sl}_\infty(K)$, if and only if p can be written as a sum of commutators; otherwise it is the entire set $M_\infty(K)$.*

5 Lie polynomials and non-associative generalizations

Next we consider a Lie polynomial p . We describe all the possible images of p in $M_2(K)$ and provide an example of such p whose image is the set of non-nilpotent trace zero matrices, together with 0. We provide an arithmetic criterion for this case. We also show that the standard polynomial s_k is not a Lie polynomial, for $k > 2$.

This section, which consists of two parts, studies the L'vov–Kaplansky conjecture for Lie algebras. Even the case of Lie identities has room for further investigation, although it has already been studied in two important books [5, 91]. In the first part we are interested in images of Lie polynomials on $M_n(K)$, viewed as a Lie algebra, and thus denoted as $\text{gl}_n(K)$ (or just gl_n if K is understood). Since $[f, g]$ can be interpreted as $fg - gf$ in the free associative algebra, we identify any Lie polynomial with an associative polynomial. In this way, any set that can arise as the image of a Lie polynomial on the Lie algebra gl_n also fits into the framework of the associative theory of $M_n(K)$, and our challenge here is to find examples of Lie polynomials that achieve the sets described in Sections 1.3.1 and 3.

As we shall see, this task is not so easy as it may seem at first glance. At the outset, the situation for Lie polynomials is subtler than for regular polynomials, for the simple reason that the most prominent polynomials in the theory, the standard polynomial s_n and the Capelli polynomial c_n , turn out not to be Lie polynomials. We first consider Lie identities, proving that the standard polynomial s_k is not a Lie polynomial for $k > 2$. A key role is played by sl_n , the Lie algebra of $n \times n$ matrices over K having trace 0.

Then we classify the possible images of Lie polynomials evaluated on 2×2 matrices, based on Section 2.1.1 where the field K was required to be quadratically closed, and Section 2.1.3, where results were provided over real closed fields, some of them holding more generally over arbitrary fields. We also consider the 3×3 case (Section 3).

5.1 Homogeneous Lie polynomials on gl_n and sl_n

We refine Conjecture 1.4 to Lie polynomials, and ask:

Question 5.1. *What is the possible image set $\text{Im } f$ of a Lie polynomial f on gl_n and sl_n ? For which Lie polynomials f do we achieve this image set? For example, what are the Lie identities of smallest degree on gl_n and sl_n ?*

In order to pass to the associative theory, we make use of the *adjoint algebra* $\text{ad } L = \{\text{ad}_a : L \rightarrow L : a \in L\}$ given by $\text{ad}_a(b) = [a, b]$. Note that

$$\dim_K(\text{ad } L) < \dim_K \text{End}_K(L) = (\dim_K L)^2.$$

The map $a \mapsto \text{ad}_a$ defines a well-known Lie homomorphism $L \rightarrow \text{ad } L$.

We write $[a_1, \dots, a_t]$ for $[a_1, \dots, [a_{t-1}, a_t]]$, and $[a^{(k)}, a_t]$ for $[a, \dots, a, a_t]$ where a occurs k times. By *ad-monomial* we mean a term $\alpha \text{ad } x_{i_1} \cdots \text{ad } x_{i_t}$ for some $\alpha \in K$. By *ad-polynomial* we mean a sum of ad-monomials.

Remark 5.2.

$$\text{ad}_{a_1} \cdots \text{ad}_{a_t}(a) = [a_1, \dots, a_t, a].$$

In this way, any ad-monomial corresponds to a Lie monomial, and thus any ad-polynomial $f(\text{ad}_{x_1}, \dots, \text{ad}_{x_t})$ gives rise to a Lie polynomial $f(x_1, \dots, x_t, y)$ taking on the same values, and in which y appears of degree 1 in each Lie monomial in the innermost set of Lie brackets.

Recall that an associative polynomial $f(x_1, \dots, x_k)$ is *alternating in the last $m+1$ variables* if f becomes 0 whenever two of the last $m+1$ variables are specialized to the same quantity. This yields:

Proposition 5.3. *Suppose L is a Lie algebra of dimension m , and $f(x_1, \dots, x_k)$ is a multilinear polynomial alternating in the last $m+1$ variables. Then*

$$f(\text{ad}_{x_1}, \dots, \text{ad}_{x_k})(y)$$

corresponds to a Lie identity of L of degree $\deg f + 1$.

Proof. The alternating property implies $f(x_1, \dots, x_k)$ vanishes on $\text{ad } L$, cf. [92, Proposition 1.2.24], so every substitution of $f(\text{ad}_{x_1}, \dots, \text{ad}_{x_k})(y)$ vanishes. ■

Since the alternating polynomial of smallest degree is the standard polynomial s_{m+1} , we have a Lie identity of degree $m+2$ for any Lie algebra of dimension m . In particular, $\dim(\text{sl}_n) = n^2 - 1$, yielding:

Corollary 5.4. *sl_n satisfies a Lie identity of degree $n^2 + 1$.*

Corollary 5.4 gives rise to the following question:

Question 5.5. *What is the minimal degree m_n of a Lie identity of sl_n ?*

By Corollary 5.4, $m_n \leq n^2 + 1$, and in particular $m_2 \leq 5$. Even the answer $m_2 = 5$ given in [91, Theorem 36.1], is not easy, although a reasonably fast combinatoric approach is given in [5, p. 165], where it is observed that any since any Lie algebra L satisfying a Lie identity of degree < 5 is solvable, one must have $m_2 \geq 5$, yielding $m_2 = 5$. Špenko [98, Proposition 7.5] looked at this from the other direction and showed that if p is a Lie polynomial of degree ≤ 4 then $\text{Im } p = \text{sl}_2$.

Conversely, we have:

Proposition 5.6. *Suppose $f(x_1, \dots, x_t, y)$ is a Lie polynomial in which y appears in degree 1 in each of its Lie monomials. Then f corresponds to an ad-polynomial taking on the same values on L as f .*

Proof. In view of Remark 5.2, it suffices to show that any Lie monomial h can be rewritten in the free Lie algebra as a sum of Lie monomials in which y appears (in degree 1) in the innermost set of Lie brackets. This could be done directly by means of the Jacobi identity, but here is a slicker argument.

Write $h = [h_1, h_2]$, and we appeal to induction on the degree of h . y appears say in h_2 . If $h_1 = y$ then we are done since $h = -[h_2, y]$ corresponds to $-\text{ad}_{h_2}$. Likewise if $h_2 = y$. In general, by induction, h_1 corresponds to some ad-monomial $\text{ad}_{x_{i_1}} \cdots \text{ad}_{x_{i_k}}(y)$ and h_2 corresponds to some ad-monomial $\text{ad}_{x_{i_{k+1}}} \cdots \text{ad}_{x_{i_\ell}}(y)$, so $[h_1, h_2]$ corresponds to $\text{ad}_{x_{i_1}} \cdots \text{ad}_{x_{i_k}}((\text{ad}_{x_{i_{k+1}}} \cdots \text{ad}_{x_{i_\ell}})(y)) = \text{ad}_{x_{i_1}} \cdots \text{ad}_{x_{i_\ell}}(y)$ as desired. ■

Corollary 5.7. *Any homogeneous Lie polynomial of degree ≥ 3 must be an identity (viewing the Lie commutator $[a, b]$ as $ab - ba$) of the Grassmann algebra G .*

Proof. Each term includes $[x_i, x_j, x_k]$, which is well known to be an identity of G . ■

Example 5.8.

- (i) The standard polynomial s_2 itself is a Lie polynomial.
- (ii) s_4 vanishes on \mathfrak{sl}_2 (viewed inside the associative algebra $M_2(K)$), since \mathfrak{sl}_2 has dimension 3. But surprisingly, this is not the polynomial of lowest degree vanishing on \mathfrak{sl}_2 , as we see next.
- (iii) Bakhturin [5, Theorem 5.14] points out that $f = [(x_1x_2 + x_2x_1), x_3]$ vanishes on \mathfrak{sl}_2 . In other words, $a_1a_2 + a_2a_1$ is scalar for any 2×2 matrices a_1, a_2 of trace 0. Indeed, a_i^2 is scalar for $i = 1, 2$, implying $a_1a_2 + a_2a_1$ is scalar unless a_1, a_2 are linearly independent, in which case $a_1a_2 + a_2a_1$ commutes with both a_1 and a_2 , and thus again is scalar. But f is not a Lie polynomial, as seen via the next lemma.

This discussion motivates us to ask when a polynomial is a Lie polynomial. Here is a very easy criterion which is of some use.

Lemma 5.9. *Any Lie polynomial which vanishes on \mathfrak{sl}_n is an identity of \mathfrak{gl}_n .*

Proof. Immediate, since $\mathfrak{gl}'_n = \mathfrak{sl}'_n$. ■

The standard polynomial s_4 is not a Lie polynomial. Here are three ways of seeing this basic fact.

- (i) Confront Example 5.8 with the fact that $m_2 = 5$, whereas $\deg s_2 = 4$.
- (ii) A computational approach. We have 15 multilinear Lie monomials of degree 4, namely $\frac{1}{2}\binom{4}{2} = 3$ of the form

$$[[x_{i_1}, x_{i_2}], [x_{i_3}, x_{i_4}]] \tag{5.1}$$

and $2\binom{4}{2} = 12$ of the form

$$[[[x_{i_1}, x_{i_2}], x_{i_3}], x_{i_4}]. \tag{5.2}$$

But

$$\begin{aligned} [[x_{i_1}, x_{i_2}], [x_{i_3}, x_{i_4}]] &= \text{ad}_{[x_{i_3}, x_{i_4}]} \text{ad}_{x_{i_2}}(x_{i_1}) \\ &= \text{ad}_{x_{i_3}} \text{ad}_{x_{i_4}} \text{ad}_{x_{i_2}}(x_{i_1}) - \text{ad}_{x_{i_4}} \text{ad}_{x_{i_3}} \text{ad}_{x_{i_2}}(x_{i_1}), \end{aligned}$$

so we can rewrite the equations (5.1) in terms of (5.2). Furthermore, with the help of the Jacobi identity, (5.2) can be reduced to seven independent Lie monomials, and one can show that these do not span s_4 . Even though this might seem unduly complicated, it provides a general program to verify that a given polynomial is not Lie.

- (iii) The third approach is simpler and works for s_k , for any $k > 2$.

P.M. Cohn was the first to tie the standard polynomial to the infinite dimensional Grassmann algebra G with base e_1, e_2, \dots , by noting that $s_k(e_1, \dots, e_k) = k!e_1 \cdots e_k \neq 0$ when $k! \neq 0$.

Theorem 5.10. *The standard polynomial s_k is not a Lie polynomial, for any $k > 2$.*

Proof. Otherwise, by Corollary 5.7 it would be an identity of G , contradicting Cohn's observation (taking $\text{Char } K = 0$). ■

5.2 The case $n = 2$

Recall from Corollary 5.4 that there is a Lie identity of degree 5.

Theorem 5.11. *If p is a homogeneous Lie polynomial evaluated on the matrix ring $M_2(K)$, where K is an algebraically closed field, then $\text{Im } f$ is either $\{0\}$, or K (the set of scalar matrices), or the set of all non-nilpotent matrices having trace zero, or $\text{sl}_2(K)$, or $M_2(K)$.*

Remark 5.12. Nonzero scalar matrices can be obtained in Theorem 5.11 only when $\text{Char } K = 2$, and the last case $M_2(K)$ is possible only if $\deg f = 1$.

Proof of Theorem 5.11. According to Theorem 2.3 the image of p must be either $\{0\}$, or K , or the set of all non-nilpotent matrices having trace zero, or $\text{sl}_2(K)$, or a dense subset of $M_2(K)$ (with respect to Zariski topology). Note that if at least one matrix having nonzero trace belongs to the image of p , then $\deg p = 1$ and thus $\text{Im } p = M_2(K)$. ■

Theorem 5.13. *For any algebraically closed field K of characteristic $\neq 2$, the image of any Lie polynomial p (not necessarily homogeneous) evaluated on $\text{sl}_2(K)$ is either $\text{sl}_2(K)$, or $\{0\}$, or the set of trace zero non-nilpotent matrices.*

Proof. For p not a PI, we can write $p = f_j + f_{j+1} + \dots + f_d$, where each f_i is a homogeneous Lie polynomial of degree i , and f_d is not PI. Therefore for any $c \in K$ we have

$$p(cx_1, cx_2, \dots, cx_m) = c^j f_j + \dots + c^d f_d.$$

Since f_d is not PI, we can take specializations of x_1, \dots, x_m for which $\det(f_d) \neq 0$. Fixing these specializations, we consider $\det(c^j f_j + \dots + c^d f_d)$ as a polynomial in c of degree $j + \dots + d$. Since the leading coefficient is not zero and K is algebraically closed, its image is K . Thus for any $k \in K$ there exist x_1, \dots, x_m for which $\det(f) = k$. Hence (for $\text{Char } K \neq 2$) any matrix with nonzero eigenvalues λ and $-\lambda$ belongs to $\text{Im } f$. Therefore $\text{Im } f$ is either sl_2 or the set of trace zero non-nilpotent matrices. ■

Let us give examples of Lie polynomials having such images:

Example 5.14 (Alexei Kanel-Belov). If $\text{Char } K = 2$, then $\text{Im } p = K$ also is possible: We take

$$p(x, y, z, t) = [[x, y], [z, t]].$$

Any value of p is the Lie product of two trace zero matrices $s_1 = [x, y]$ and $s_2 = [z, t]$. Both can be written as $s_i = h_i + u_i + v_i$, where the h_i are diagonal trace zero matrices (which are scalar since $\text{Char } K = 2$), the u_i are proportional to e_{12} , and the v_i are proportional to e_{21} . Thus $[s_1, s_2] = [u_1, v_2] + [u_2, v_1]$ is scalar.

Over an arbitrary field, $\text{Im } p$ can indeed be equal to $\{0\}$, or K , or the set of all non-nilpotent matrices having trace zero, or $\text{sl}_2(K)$, or $M_2(K)$.

- (i) $\text{Im } x = M_2(K)$.
- (ii) $\text{Im}[x, y] = \text{sl}_2$.
- (iii) Next, we construct a Lie polynomial whose image evaluated on $\text{sl}_2(K)$ is the set of all non-nilpotent matrices having trace zero. We take the multilinear polynomial $h(u_1, \dots, u_8)$ constructed in [33] by Drensky and Kasparian which is central on 3×3 matrices. Given 2×2 matrices x_1, \dots, x_9 we consider the homogeneous Lie polynomial

$$p(x_1, \dots, x_9) = h(\text{ad}_{[x_9, x_9, \dots, x_9, x_1]}, \text{ad}_{x_2}, \text{ad}_{x_3}, \dots, \text{ad}_{x_8})(x_9).$$

For any 2×2 matrix x , ad_x is a 3×3 matrix since \mathfrak{sl}_2 is 3-dimensional; hence, for any values of x_i , the value of p has to be proportional to x_9 . However for x_9 nilpotent, this must be zero, since $[x^{(3)}, y] = 0$ for any $y \in \mathfrak{sl}_2(K)$ if x is nilpotent. (When we open the brackets we have the sum of 8 terms and each term equals $x^k y x^{3-k}$. But for any integer k , either $k \geq 2$ or $3 - k \geq 2$.) Thus the image of p is exactly the set of non-nilpotent trace zero matrices.

Another example of a homogeneous Lie polynomial with no nilpotent values is $p(x, y) = [[x, y], x], [[x, y], y]$. (See [8, Example 4.9] for details.)

5.3 The case $n = 3$

New questions arise concerning the possible evaluation of Lie polynomials on $M_n(K)$.

According to Theorem 1.18, if p is a homogeneous polynomial with trace vanishing image, then $\text{Im } p$ is one of the following:

- $\{0\}$,
- the set of scalar matrices (which can occur only if $\text{Char } K = 3$),
- a dense subset of $\mathfrak{sl}_3(K)$, or
- the set of 3-scalar matrices, i.e., with eigenvalues $(c, c\omega, c\omega^2)$, where ω is our cube root of 1.

Drensky and Rashkova [35] have found several identities of \mathfrak{sl}_3 of degree 6, but they cannot be Lie polynomials, since otherwise they would be identities of \mathfrak{gl}_3 and thus a multiple of s_6 , which is not a Lie polynomial. Thus, one must go to higher degree.

In the associative case, the fact that the generic division algebra has a 3-central element implies that there is a homogeneous 3-central polynomial for $M_3(K)$, i.e., all of whose values take on eigenvalues $c, \omega c, \omega^2 c$, where ω is a cube root of 1. But any matrix with these eigenvalues is either scalar or has trace 0. This leads us to the basic questions needed to complete the case $n = 3$:

Question 5.15. *Is there a Lie polynomial p whose values on $\mathfrak{sl}_3(\mathbb{C})$ are dense, but does not take on all values?*

Question 5.16. *Is there a Lie polynomial p whose values on \mathfrak{sl}_3 all take on eigenvalues $c, \omega c, \omega^2 c$, where ω is a primitive cube root of 1?*

5.3.1 A group theoretical question and its relation to the Lie theoretical problem

There is a group conjecture (see [57, Question 2] for the more general case):

Conjecture 5.17. *If the field K is algebraically closed of characteristic 0, then the image of any nontrivial group word $w(x_1, \dots, x_m)$ on the projective linear group $\text{PSL}_2(K)$ is $\text{PSL}_2(K)$.*

Remark 5.18. Note that if one takes the group SL_2 instead of PSL_2 , Conjecture 5.17 fails, since the matrix $-I + e_{12}$ does not belong to the image of the word map $w = x^2$.

Example 5.19. When $\text{Char } K = p > 0$, the image of the word map $w(x) = x^p$ evaluated on $\text{PSL}_2(K)$ is not $\text{PSL}_2(K)$. Indeed, otherwise the matrix $I + e_{12}$ could be written as x^p for $x \in \text{PSL}_2(K)$. If the eigenvalues of x are equal, then $x = I + n$ where n is nilpotent. Therefore $x^p = (I + n)^p = I + pn = I$. If the eigenvalues of x are not equal, then x is diagonalizable and therefore x^p is also diagonalizable, a contradiction.

Lemma 5.20 (Liebeck, Nikolov, Shalev, cf. also [10, 44]). *Im w contains all matrices from $\mathrm{PSL}_2(K)$ which are not unipotent.*

Proof. According to [16] the image of the word map w must be Zariski dense in $\mathrm{SL}_2(K)$. Therefore the image of $\mathrm{tr} w$ must be Zariski dense in K . Note that $\mathrm{tr} w$ is a homogeneous rational function and K is algebraically closed. Hence, $\mathrm{Im}(\mathrm{tr} w) = K$. For any $\lambda \neq \pm 1$ any matrix with eigenvalues λ and λ^{-1} belongs to the image of w since there is a matrix with trace $\lambda + \lambda^{-1}$ in $\mathrm{Im} w$ and any two matrices from SL_2 with equal trace (except trace ± 2) are similar. Note that the identity matrix I belongs to the image of any word map. ■

However the question whether one of the matrices $(I + e_{12})$ or $(-I - e_{12})$ (which are equal in PSL_2) belongs to the image of w remains open. We conjecture that $I + e_{12}$ must belong to $\mathrm{Im} w$. Note that if there exists i such that the degree of x_i in w is $k \neq 0$ then we can consider all $x_j = I$ for $j \neq i$ and $x_i = I + e_{12}$. Then the value of w is $(I + e_{12})^k = I + ke_{12}$ and this is a unipotent matrix since $\mathrm{Char} K = 0$, and thus $\mathrm{Im} w = \mathrm{PSL}_2(K)$. Therefore it is interesting to consider word maps $w(x_1, \dots, x_m)$ such that the degree of each x_i is zero.

This is why Conjecture 5.17 can be reformulated as follows:

Conjecture 5.21. *Let $w(x_1, \dots, x_m)$ be a group word whose degree at each x_i is 0. Then the image of w on $G = \mathrm{GL}_2(K)/\{\pm 1\}$ must be $\mathrm{PSL}_2(K)$.*

One can consider matrices $z_i = \frac{x_i}{\sqrt{\det x_i}}$ and note that $w(z_1, \dots, z_m) = w(x_1, \dots, x_m)$.

For Conjecture 5.21 we take $y_i = x_i - I$. Then we can open the brackets in

$$w(1 + y_1, 1 + y_2, \dots, 1 + y_m) = 1 + f(y_1, \dots, y_m) + g(y_1, \dots, y_m),$$

where f is a homogeneous Lie polynomial of degree d , and g is the sum of terms of degree greater than d . Therefore it is interesting to investigate the possible images of Lie polynomials, whether it is possible that the image of l does not contain nilpotent matrices. Unfortunately we saw such an example (Example 5.14, although its degree must be at least 5 by Špenko [103, Lemma 7.4]). More general questions about surjectivity of word maps in groups and polynomials in algebras are considered in [57].

Remark 5.22. Our next theorem describes the situation in which the trace vanishing polynomial does not take on nonzero nilpotent values. It implies that any nontrivial word map w evaluated on PSL_2 is not surjective iff its projection to \mathfrak{sl}_2 given by $\mathfrak{sl}_2: x \mapsto x - \frac{1}{2} \mathrm{tr} x$ is a multiple of any prime divisor of $\det(\pi(w))$. This might help in answering Conjecture 5.17.

Theorem 5.23. *Let $p(x_1, \dots, x_m)$ be a trace vanishing polynomial, evaluated on $M_n(K[\xi])$. Let $\bar{p} = p(y_1, \dots, y_m)$. Then p takes on no nonzero nilpotent values on any integral domain containing K , iff each prime divisor d of $\det(\bar{p})$ also divides each entry of \bar{p} .*

Proof. (\Rightarrow) If some prime divisor d of $\det(\bar{p})$ does not divide \bar{p} , then \bar{p} does not specialize to 0 modulo d . Therefore we have a nonzero matrix in the image of p which has determinant zero and also trace zero, and thus is nilpotent, a contradiction.

(\Leftarrow) Assume that p takes on a nonzero nilpotent value over some integral domain extension of K . Thus $\det \bar{p}$ goes to 0 under the corresponding specialization of the $\xi_{i,j}^k$, so some prime divisor d of $\det(\bar{p})$ goes to 0, and \bar{p} is not divisible by d . ■

5.4 Open problems for nonassociative algebras

We described the evaluations of Lie polynomials on \mathfrak{sl}_2 and \mathfrak{gl}_2 . One can consider the same problem for other important Lie algebras:

Problem 5.24. *Give a classification of possible images of multilinear/homogeneous/arbitrary Lie polynomials on the Lie algebra*

- \mathfrak{so}_4 (an algebra of rank 2),
- \mathfrak{so}_6 (an algebra of rank 3),
- \mathfrak{sp}_4 (an algebra of rank 2),
- \mathfrak{sp}_6 (an algebra of rank 3),
- G_2 (an algebra of rank 2),
- F_4 (an algebra of rank 4).

Analogous problems occur for the non-Lie case:

Problem 5.25. *Investigate possible evaluations of multilinear/homogeneous/arbitrary polynomials on simple Jordan algebras.*

Problem 5.26. *Investigate possible evaluations of multilinear/homogeneous/arbitrary polynomials on the Malcev algebra (for example, see [109]).*

Problem 5.27. *There are nonassociative operator algebras (in particular, the algebra of left multiplications). It is important to investigate possible evaluations of polynomials on these algebras.*

Acknowledgements

We would like to thank M. Brešar, B. Kunyavskii and express our special gratitude to E. Plotkin for interesting and fruitful discussions. We would also like to thank the anonymous referees whose suggestions contributed significantly to the level of the paper. The second and third named authors were supported by the ISF (Israel Science Foundation) grant 1994/20. The first named author was supported by the Russian Science Foundation grant No. 17-11-01377. The second and fourth named authors were supported by Israel Innovation Authority, grant no. 63412: Development of A.I. based platform for e commerce.

References

- [1] Albert A.A., Muckenhoupt B., On matrices of trace zeros, *Michigan Math. J.* **4** (1957), 1–3.
- [2] Almutairi N.B., On multilinear polynomials evaluated on quaternion algebra, Master Thesis, Kent State University, 2016.
- [3] Amitsur A.S., Levitzki J., Minimal identities for algebras, *Proc. Amer. Math. Soc.* **1** (1950), 449–463.
- [4] Anzis B.E., Emrich Z.M., Valiveti K.G., On the images of Lie polynomials evaluated on Lie algebras, *Linear Algebra Appl.* **469** (2015), 51–75.
- [5] Bahturin Yu.A., Identical relations in Lie algebras, VNU Science Press, Utrecht, 1987.
- [6] Bandman T., Garion S., Grunewald F., On the surjectivity of Engel words on $\mathrm{PSL}(2, q)$, *Groups Geom. Dyn.* **6** (2012), 409–439, [arXiv:1008.1397](https://arxiv.org/abs/1008.1397).
- [7] Bandman T., Garion S., Kunyavskii B., Equations in simple matrix groups: algebra, geometry, arithmetic, dynamics, *Cent. Eur. J. Math.* **12** (2014), 175–211, [arXiv:1302.4667](https://arxiv.org/abs/1302.4667).
- [8] Bandman T., Gordeev N., Kunyavskii B., Plotkin E., Equations in simple Lie algebras, *J. Algebra* **355** (2012), 67–79, [arXiv:1012.4106](https://arxiv.org/abs/1012.4106).
- [9] Bandman T., Kunyavskii B., Criteria for equidistribution of solutions of word equations on $\mathrm{SL}(2)$, *J. Algebra* **382** (2013), 282–302, [arXiv:1201.5260](https://arxiv.org/abs/1201.5260).
- [10] Bandman T., Zarhin Y.G., Surjectivity of certain word maps on $\mathrm{PSL}(2, \mathbb{C})$ and $\mathrm{SL}(2, \mathbb{C})$, *Eur. J. Math.* **2** (2016), 614–643, [arXiv:1407.3447](https://arxiv.org/abs/1407.3447).

- [11] Belov A.Ya., Counterexamples to the Specht problem, *Sb. Math.* **191** (2000), 329–340.
- [12] Belov A.Ya., Borisenko V.V., Latyshev V.N., Monomial algebras, *J. Math. Sci.* **87** (1997), 3463–3575.
- [13] Bokut L.A., Unsolvability of the equality problem and subalgebras of finitely presented Lie algebras, *Math. USSR-Izv.* **6** (1972), 1153–1199.
- [14] Bokut L.A., Chen Y., Gröbner–Shirshov bases and their calculation, *Bull. Math. Sci.* **4** (2014), 325–395, [arXiv:1303.5366](#).
- [15] Bokut L.A., Kukin G.P., Undecidable algorithmic problems for semigroups, groups and rings, *J. Soviet Math.* **45** (1989), 871–911.
- [16] Borel A., On free subgroups of semisimple groups, *Enseign. Math.* **29** (1983), 151–164.
- [17] Brešar M., Commutators and images of noncommutative polynomials, [arXiv:2001.10392](#).
- [18] Brešar M., Klep I., Values of noncommutative polynomials, Lie skew-ideals and tracial Nullstellensätze, *Math. Res. Lett.* **16** (2009), 605–626, [arXiv:0810.1774](#).
- [19] Brešar M., Procesi C., Špenko Š., Quasi-identities on matrices and the Cayley–Hamilton polynomial, *Adv. Math.* **280** (2015), 439–471, [arXiv:1212.4597](#).
- [20] Burness T.C., Liebeck M.W., Shalev A., The length and depth of algebraic groups, *Math. Z.* **291** (2019), 741–760, [arXiv:1712.08214](#).
- [21] Burness T.C., Liebeck M.W., Shalev A., The length and depth of compact Lie groups, *Math. Z.* **294** (2020), 1457–1476, [arXiv:1805.09893](#).
- [22] Buzinski D., Winstanley R., On multilinear polynomials in four variables evaluated on matrices, *Linear Algebra Appl.* **439** (2013), 2712–2719, [arXiv:1511.06331](#).
- [23] Chuang C.-L., On ranges of polynomials in finite matrix rings, *Proc. Amer. Math. Soc.* **110** (1990), 293–302.
- [24] Cohn P.M., The range of derivations on a skew field and the equation $ax - xb = c$, *J. Indian Math. Soc. (N.S.)* **37** (1973), 61–69.
- [25] Cox D., Little J., O’Shea D., Ideals, varieties, and algorithms. An introduction to computational algebraic geometry and commutative algebra, 3rd ed., *Undergraduate Texts in Mathematics*, Springer, New York, 2007.
- [26] De Concini C., Papi P., Procesi C., The adjoint representation inside the exterior algebra of a simple Lie algebra, *Adv. Math.* **280** (2015), 21–46, [arXiv:1311.4338](#).
- [27] De Concini C., Procesi C., The invariant theory of matrices, *University Lecture Series*, Vol. 69, Amer. Math. Soc., Providence, RI, 2017.
- [28] Deligne P., Sullivan D., Division algebras and the Hausdorff–Banach–Tarski paradox, *Enseign. Math.* **29** (1983), 145–150.
- [29] Dniester Notebook: Unsolved problems in the theory of rings and modules, 4th ed., Mathematics Institute, Russian Academy of Sciences Siberian Branch, Novosibirsk, 1993.
- [30] Donkin S., Invariants of several matrices, *Invent. Math.* **110** (1992), 389–401.
- [31] Drensky V., Free algebras and PI-algebras. Graduate course in algebra, Springer-Verlag Singapore, Singapore, 2000.
- [32] Drensky V., Formanek E., Polynomial identity rings, *Advanced Courses in Mathematics. CRM Barcelona*, Birkhäuser Verlag, Basel, 2004.
- [33] Drensky V., Kasparyan A., A new central polynomial for 3×3 matrices, *Comm. Algebra* **13** (1985), 745–752.
- [34] Drensky V., Piacentini Cattaneo G.M., A central polynomial of low degree for 4×4 matrices, *J. Algebra* **168** (1994), 469–478.
- [35] Drensky V., Rashkova T.G., Weak polynomial identities for the matrix algebras, *Comm. Algebra* **21** (1993), 3779–3795.
- [36] Dykema K.J., Klep I., Instances of the Kaplansky–Lvov multilinear conjecture for polynomials of degree three, *Linear Algebra Appl.* **508** (2016), 272–288, [arXiv:1508.01238](#).
- [37] Egorchenkova E., Gordeev N., Products of three word maps on simple algebraic groups, *Arch. Math. (Basel)* **112** (2019), 113–122.
- [38] Ellers E.W., Gordeev N., On the conjectures of J. Thompson and O. Ore, *Trans. Amer. Math. Soc.* **350** (1998), 3657–3671.

- [39] Fagundes P.S., The images of multilinear polynomials on strictly upper triangular matrices, *Linear Algebra Appl.* **563** (2019), 287–301, [arXiv:1807.09136](#).
- [40] Fagundes P.S., de Mello T.C., Images of multilinear polynomials of degree up to four on upper triangular matrices, *Oper. Matrices* **13** (2019), 283–292, [arXiv:1807.09421](#).
- [41] Formanek E., Central polynomials for matrix rings, *J. Algebra* **23** (1972), 129–132.
- [42] Formanek E., The polynomial identities and invariants of $n \times n$ matrices, *CBMS Regional Conference Series in Mathematics*, Vol. 78, Amer. Math. Soc., Providence, RI, 1991.
- [43] Giambruno A., Zaicev M., Polynomial identities and asymptotic methods, *Mathematical Surveys and Monographs*, Vol. 122, Amer. Math. Soc., Providence, RI, 2005.
- [44] Gordeev N., On Engel words on simple algebraic groups, *J. Algebra* **425** (2015), 215–244.
- [45] Gordeev N., Kunyavskii B., Plotkin E., Geometry of word equations in simple algebraic groups over special fields, *Russian Math. Surveys* **73** (2018), 753–796, [arXiv:1808.02303](#).
- [46] Gordeev N., Kunyavskii B., Plotkin E., Word maps on perfect algebraic groups, *Internat. J. Algebra Comput.* **28** (2018), 1487–1515, [arXiv:1801.00381](#).
- [47] Gordeev N., Kunyavskii B., Plotkin E., Word maps, word maps with constants and representation varieties of one-relator groups, *J. Algebra* **500** (2018), 390–424, [arXiv:1801.00379](#).
- [48] Gordienko A.S., Lecture notes on polynomial identities, Preprint, 2014.
- [49] Guterman A.E., Kuzma B., Maps preserving zeros of matrix polynomials, *Dokl. Math.* **80** (2009), 508–510.
- [50] Halpin P., Central and weak identities for matrices, *Comm. Algebra* **11** (1983), 2237–2248.
- [51] Helling H., Eine Kennzeichnung von Charakteren auf Gruppen und assoziativen Algebren, *Comm. Algebra* **1** (1974), 491–501.
- [52] Herstein I.N., On the Lie structure of an associative ring, *J. Algebra* **14** (1970), 561–571.
- [53] Hui C.Y., Larsen M., Shalev A., The Waring problem for Lie groups and Chevalley groups, *Israel J. Math.* **210** (2015), 81–100, [arXiv:1404.4786](#).
- [54] Itoh M., Invariant theory in exterior algebras and Amitsur–Levitzi type theorems, *Adv. Math.* **288** (2016), 679–701, [arXiv:1404.1980](#).
- [55] Kanel-Belov A., Grigoriev S., Elishev A., Yu J.-T., Zhang W., Lifting of polynomial symplectomorphisms and deformation quantization, *Comm. Algebra* **46** (2018), 3926–3938, [arXiv:1707.06450](#).
- [56] Kanel-Belov A., Karasik Y., Rowen L.H., Computational aspects of polynomial identities, Vol. 1, Kemer’s theorems, 2nd ed., *Monographs and Research Notes in Mathematics*, CRC Press, Boca Raton, FL, 2016.
- [57] Kanel-Belov A., Kunyavskii B., Plotkin E., Word equations in simple groups and polynomial equations in simple algebras, *Vestnik St. Petersburg Univ. Math.* **46** (2013), 3–13, [arXiv:1304.5052](#).
- [58] Kanel-Belov A., Malev S., Rowen L., The images of non-commutative polynomials evaluated on 2×2 matrices, *Proc. Amer. Math. Soc.* **140** (2012), 465–478, [arXiv:1005.0191](#).
- [59] Kanel-Belov A., Malev S., Rowen L., The images of multilinear polynomials evaluated on 3×3 matrices, *Proc. Amer. Math. Soc.* **144** (2016), 7–19, [arXiv:1306.4389](#).
- [60] Kanel-Belov A., Malev S., Rowen L., Power-central polynomials on matrices, *J. Pure Appl. Algebra* **220** (2016), 2164–2176, [arXiv:1310.1598](#).
- [61] Kanel-Belov A., Malev S., Rowen L., The images of Lie polynomials evaluated on matrices, *Comm. Algebra* **45** (2017), 4801–4808, [arXiv:1506.06792](#).
- [62] Kanel-Belov A., Rowen L.H., Computational aspects of polynomial identities, *Research Notes in Mathematics*, Vol. 9, A K Peters, Ltd., Wellesley, MA, 2005.
- [63] Kemer A.R., Capelli identities and nilpotence of the radical of a finitely generated PI-algebra, *Sov. Math. Dokl.* **255** (1980), 750–753.
- [64] Kemer A.R., Varieties and Z_2 -graded algebras, *Math. USSR Izv.* **25** (1985), 359–374.
- [65] Klimenko E., Kunyavskii B., Morita J., Plotkin E., Word maps in Kac–Moody setting, *Toyama Math. J.* **37** (2015), 25–54, [arXiv:1506.01422](#).
- [66] Kolesnikov P.S., The Makar-Limanov algebraically closed skew field, *Algebra Logic* **39** (2000), 378–395.
- [67] Kolesnikov P.S., On various definitions of algebraically closed skew fields, *Algebra Logic* **40** (2001), 219–230.
- [68] Kulyamin V.V., Images of graded polynomials in matrix rings over finite group algebras, *Russian Math. Surveys* **55** (2000), 345–346.

- [69] Kulyamin V.V., On images of polynomials in finite matrix rings, Ph.D. Thesis, Moscow Lomonosov State University, Moscow, 2000.
- [70] Larsen M., Word maps have large image, *Israel J. Math.* **139** (2004), 149–156, [arXiv:math.GR/0211302](#).
- [71] Larsen M., Shalev A., Word maps and Waring type problems, *J. Amer. Math. Soc.* **22** (2009), 437–466, [arXiv:math.GR/0701334](#).
- [72] Larsen M., Shalev A., Tiep P.H., The Waring problem for finite simple groups, *Ann. of Math.* **174** (2011), 1885–1950.
- [73] Larsen M., Shalev A., Tiep P.H., Waring problem for finite quasisimple groups, *Int. Math. Res. Not.* **2013** (2013), 2323–2348, [arXiv:1107.3341](#).
- [74] Larsen M., Tiep P.H., A refined Waring problem for finite simple groups, *Forum Math. Sigma* **3** (2015), e6, 22 pages, [arXiv:1312.4998](#).
- [75] Lee T.-K., Zhou Y., Right ideals generated by an idempotent of finite rank, *Linear Algebra Appl.* **431** (2009), 2118–2126.
- [76] Li C., Tsui M.C., On the images of multilinear maps of matrices over finite-dimensional division algebras, *Linear Algebra Appl.* **493** (2016), 399–410.
- [77] Liebeck M.W., O’Brien E.A., Shalev A., Tiep P.H., The Ore conjecture, *J. Eur. Math. Soc. (JEMS)* **12** (2010), 939–1008.
- [78] Ma A., Oliva J., On the images of Jordan polynomials evaluated over symmetric matrices, *Linear Algebra Appl.* **492** (2016), 13–25.
- [79] Makar-Limanov L., Algebraically closed skew fields, *J. Algebra* **93** (1985), 117–135.
- [80] Makar-Limanov L., An example of a skew field without a trace, *Comm. Algebra* **17** (1989), 2303–2307.
- [81] Malev S., The images of non-commutative polynomials evaluated on 2×2 matrices over an arbitrary field, *J. Algebra Appl.* **13** (2014), 1450004, 12 pages, [arXiv:1310.8563](#).
- [82] Malev S., The images of noncommutative polynomials evaluated on the quaternion algebra, *J. Algebra Appl.*, to appear, [arXiv:1906.04973](#).
- [83] Malev S., Pines C., The images of multilinear non-associative polynomials evaluated on a rock-paper-scissors algebra with unit over an arbitrary field, [arXiv:2006.04517](#).
- [84] Malle G., The proof of Ore’s conjecture (after Ellers–Gordeev and Liebeck–O’Brien–Shalev–Tiep), *Astérisque* **361** (2014), Exp. No. 1069, 325–348.
- [85] Mesyan Z., Polynomials of small degree evaluated on matrices, *Linear Multilinear Algebra* **61** (2013), 1487–1495, [arXiv:1212.1925](#).
- [86] Procesi C., The invariant theory of $n \times n$ matrices, *Adv. Math.* **19** (1976), 306–381.
- [87] Procesi C., On the theorem of Amitsur–Levitzki, *Israel J. Math.* **207** (2015), 151–154.
- [88] Razmyslov Yu.P., Finite basing of the identities of a matrix algebra of second order over a field of characteristic zero, *Algebra Logic* **12** (1973), 47–63.
- [89] Razmyslov Yu.P., On a problem of Kaplansky, *Math USSR. Izv.* **7** (1973), 479–496.
- [90] Razmyslov Yu.P., Trace identities of full matrix algebras over a field of characteristic zero, *Math USSR. Izv.* **38** (1974), 727–760.
- [91] Razmyslov Yu.P., Identities of algebras and their representations, *Translations of Mathematical Monographs*, Vol. 138, Amer. Math. Soc., Providence, RI, 1994.
- [92] Rowen L.H., Polynomial identities in ring theory, *Pure and Applied Mathematics*, Vol. 84, Academic Press, Inc., New York – London, 1980.
- [93] Rowen L.H., Graduate algebra: commutative view, *Graduate Studies in Mathematics*, Vol. 73, Amer. Math. Soc., Providence, RI, 2006.
- [94] Rowen L.H., Graduate algebra: noncommutative view, *Graduate Studies in Mathematics*, Vol. 91, Amer. Math. Soc., Providence, RI, 2008.
- [95] Saltman D.J., On p -power central polynomials, *Proc. Amer. Math. Soc.* **78** (1980), 11–13.
- [96] Santulo E.A., Yasurmura F.Y., On the image of polynomials evaluated on incidence algebras: a counterexample and a solution, [arXiv:1902.08116](#).
- [97] Shalev A., Commutators, words, conjugacy classes and character methods, *Turkish J. Math.* **31** (2007), suppl., 131–148.

-
- [98] Shalev A., Word maps, conjugacy classes, and a noncommutative Waring-type theorem, *Ann. of Math.* **170** (2009), 1383–1416.
- [99] Shalev A., Some results and problems in the theory of word maps, in Erdős Centennial, *Bolyai Soc. Math. Stud.*, Vol. 25, *János Bolyai Math. Soc.*, Budapest, 2013, 611–649.
- [100] Shirshov A.I., Some questions in the theory of rings close to associative, *Russian Math. Surveys* **13** (1958), no. 6, 3–20.
- [101] Thom A., Convergent sequences in discrete groups, *Canad. Math. Bull.* **56** (2013), 424–433, [arXiv:1003.4093](#).
- [102] Vitas D., Multilinear polynomials are surjective on algebras with surjective inner derivations, Preprint, 2020.
- [103] Špenko Š., On the image of a noncommutative polynomial, *J. Algebra* **377** (2013), 298–311, [arXiv:1212.4600](#).
- [104] Wang Y., The images of multilinear polynomials on 2×2 upper triangular matrix algebras, *Linear Multilinear Algebra* **67** (2019), 2366–2372.
- [105] Wang Y., The image of arbitrary polynomials on 2×2 upper triangular matrix algebras, Preprint, 2019.
- [106] Wang Y., Liu P.P., Bai J., Correction: “The images of multilinear polynomials on 2×2 upper triangular matrix algebras”, *Linear Multilinear Algebra* **67** (2019), no. 11, i–vi.
- [107] Watkins W., Linear maps that preserve commuting pairs of matrices, *Linear Algebra Appl.* **14** (1976), 29–35.
- [108] Zelmanov E., Infinite algebras and pro- p groups, in Infinite groups: geometric, combinatorial and dynamical aspects, *Progr. Math.*, Vol. 248, *Birkhäuser*, Basel, 2005, 403–413.
- [109] Zhevlakov K.A., Slinko A.M., Shestakov I.P., Shirshov A.I., Rings close to associative, Moscow, Nauka, 1978.
- [110] Zubkov A.N., Non-abelian free pro- p -groups cannot be represented by 2-by-2 matrices, *Sib. Math. J.* **28** (1987), 742–747.