The L'vov–Kaplansky conjecture for polynomials of degree four

Qiyue Tang

November 5, 2026

Abstract

We prove the L'vov–Kaplansky conjecture for multilinear polynomials of degree four over algebraically closed fields of characteristic zero.

hexagon:2611.00119v2

1 Introduction

The commutator problem is an early instance of the problem of describing polynomial images on matrix algebras. For matrices A, B over a field, the additive commutator $[A, B] = AB - BA$ has trace zero, since $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. Shoda proved the converse in characteristic zero: every matrix of trace zero is one commutator [1, Satz 3]. Albert and Muckenhoupt extended the result to arbitrary fields [2]. Thus a linear condition on matrix entries completely describes the image of this particular noncommutative polynomial. Replacing the commutator by a general polynomial leads to the image problem considered here.

Throughout this paper K is an algebraically closed field of characteristic zero. Let $M_n(K)$ be the algebra of $n \times n$ matrices, I_n its identity matrix, and $\operatorname{tr} T = \sum_{i=1}^{n} T_{ii}$ the ordinary, unnormalized trace. Write

$$\mathfrak{sl}_n(K) = \{T \in M_n(K) : \operatorname{tr} T = 0\}.$$

The free associative algebra $K\langle X_1, \dots, X_m\rangle$ has the words in its noncommuting generators as a K-basis, with multiplication by concatenation. An ordinary multilinear polynomial of degree m is

$$p(X_1, \dots, X_m) = \sum_{\sigma \in S_m} a_\sigma X_{\sigma(1)} \cdots X_{\sigma(m)}, \qquad a_\sigma \in K,$$

where S_m is the permutation group of $\{1, \dots, m\}$. The integer m counts inputs and the integer n is the matrix order. The image is the set

$$p(M_n(K)) = \{p(A_1, \dots, A_m) : A_i \in M_n(K)\}.$$

A polynomial identity has zero image. A central polynomial has nonzero image consisting of scalar matrices. These two possibilities arise naturally in polynomial identity theory: the standard identity of Amitsur and Levitzki [3] and the central polynomials of Formanek [4] give examples.

The L'vov–Kaplansky question asks whether $p(M_n(K))$ is always a vector space. It is recorded as Question 1.98 of the fourth edition of the Dniester Notebook [7, Introduction]; Kanel-Belov, Malev, and Rowen discuss its attribution to Kaplansky and its formulation by

2020 Mathematics Subject Classification. Primary 16R10; Secondary 15A24, 16R30.
Keywords. Multilinear polynomial, matrix algebra, L'vov–Kaplansky conjecture.

L'vov [10, Introduction]. Over the present field, an affirmative answer says that the image is one of

$$\{0\}, \qquad KI_n, \qquad \mathfrak{sl}_n(K), \qquad M_n(K).$$

Each of these possibilities occurs among multilinear polynomials. The zero polynomial and the identity input give the first and fourth, the commutator gives the third, and multilinear central polynomials give the second.

The linear span of the image has a simpler description than the image itself. Brešar and Klep proved that this span is a Lie ideal and classified it by the polynomial's identities and trace [5, Theorems 2.3 and 4.5]. For a multilinear polynomial, the underlying identity is

$$[p(A_1, \ldots, A_m), T] = \sum_{i=1}^{m} p(A_1, \ldots, [A_i, T], \ldots, A_m).$$

Their tracial Nullstellensatz identifies polynomials of zero trace with sums of commutators modulo matrix identities. The image problem asks which matrices occur as single values. In particular, passing from traceless containment to full image requires attaining each conjugacy class of nonzero trace.

There are two useful ways to restrict the question: fix the matrix order and allow arbitrary polynomial degree, or fix the degree and allow arbitrary matrix order. Kanel-Belov, Malev, and Rowen proved the full image classification in matrix order two [10, Theorem 2]. Their work combines conjugation-invariant cones, generic matrices, and invariant theory. In matrix order three, their subsequent paper obtains image and density restrictions, with possible power-central behavior [8].

Mesyan generalized the commutator theorem by proving that every nonzero multilinear polynomial of degree at most three contains all traceless matrices in its image, in the range of field sizes stated in [6, Theorem 13]. He proposed the corresponding containment for a polynomial of degree m when $n \geq m-1$. This weaker problem is usually called the Mesyan conjecture. Buzinski and Winstanley treated degree four for algebraically closed fields of characteristic zero and $n \geq 3$ [7]. A lemma in that argument required correction. Fagundes, de Mello, and dos Santos supplied the correction and established

$$\mathfrak{sl}_n(K) \subseteq p(M_n(K)) \qquad (n \geq 3) \tag{1}$$

for every nonzero multilinear quartic [13, Theorem 8]. Thus the trace-zero quartic case follows from existing work. The trace-nonzero case requires attainment of the remaining conjugacy classes.

For the full conjecture in degree three, Dykema and Klep proved the result for every even matrix order and every odd order below 17 [11, Theorem 1.2]. Their proof uses Sylvester's clock-and-shift basis and determinant conditions on the six word coefficients. Vitas proved the degree-three conjecture in all orders [12, Theorem B], by constructing slices covering prescribed Jordan forms and solving the resulting coefficient conditions. We use his theorem when replacing one input of a quartic by the identity leaves a cubic whose trace is nonzero.

We prove the full degree-four conjecture. The additional conclusion beyond (1) is that every trace-nonzero quartic has image $M_n(K)$ for $n \geq 3$.

Theorem 1.1. *Let p be an ordinary multilinear polynomial of degree four over K. For every $n \geq 3$,*

$$p(M_n(K)) = \begin{cases} \{0\}, & p = 0, \\ \mathfrak{sl}_n(K), & p \neq 0 \text{ and } \operatorname{tr} p \text{ is identically zero}, \\ M_n(K), & \operatorname{tr} p \text{ is not identically zero}. \end{cases}$$

Consequently the image of p is a vector space in every positive matrix order.

Two words are cyclically equivalent when one is obtained by moving an initial segment

of the other to its end. The formal trace of a polynomial is its class in the vector space with cyclic word classes as a basis. For a quartic the six class coefficients are the sums of the four word coefficients in each class. Matrix trace vanishes on the difference of cyclically equivalent words. Conversely, rotate each word to place X_4 last and let $W(X_1, X_2, X_3)$ be the resulting cubic. Then

$$\operatorname{tr} p(A_1, A_2, A_3, A_4) = \operatorname{tr}\big(W(A_1, A_2, A_3)A_4\big).$$

The trace pairing is the bilinear form $(U, V) \mapsto \operatorname{tr}(UV) = \sum_{i,j} U_{ij}V_{ji}$; it is nondegenerate because each entry of U can be tested by a suitable V. If the left side vanishes identically, this pairing makes W an identity on $M_n(K)$. The minimal nonzero identity degree is $2n$ [3], so $W = 0$ for $n \geq 3$. This proves equivalence between formal and evaluated trace vanishing in Theorem 1.1.

We fix all but one input and study the image of the resulting linear map. Weighted permutation inputs produce slices containing subspaces defined by prescribed zero entries. Scalar shifts of identity reductions produce slices of rank at least $n^2 - 1$. The coefficients left by these constructions form a six-dimensional space. Two involution substitutions reduce it to eight families, and five explicit 3×3 slices settle each family. A block construction passes from sizes two and three to every larger size. The coefficient matrices, minors, and factor-by-factor elimination calculations are given in the appendices.

2 Linear slices and conjugacy classes

For a chosen input position, a *slice* is the linear map obtained by fixing the other inputs. Its image is a subspace of $M_n(K)$, and its rank is the dimension of that image. Slice rank conditions are Zariski open conditions on the frozen inputs, since they assert the nonvanishing of a minor. We repeatedly intersect finitely many nonempty open subsets of an affine space. Such an intersection is nonempty because its coordinate ring is an integral domain. Matrix indices in the proofs below run from 1 to n, unless explicitly stated otherwise. The matrix unit E_{ij} has its only nonzero entry, equal to one, in position (i, j).

A conjugacy class of T is $\{gTg^{-1} : g \in \mathrm{GL}_n(K)\}$, where $\mathrm{GL}_n(K)$ is the group of invertible matrices; we write $C \sim T$ when C belongs to this class. For a subspace H, its orthogonal space for the trace pairing is $H^{\perp} = \{A : \operatorname{tr}(AX) = 0 \text{ for every } X \in H\}$. The codimension of H is the dimension of its ambient space minus $\dim H$.

Lemma 2.1. *For $n \geq 2$, every nonzero conjugation-invariant subspace of $\mathfrak{sl}_n(K)$ is $\mathfrak{sl}_n(K)$. The span of the conjugates of a nonscalar matrix contains $\mathfrak{sl}_n(K)$. The differences of matrices in any nonscalar conjugacy class span $\mathfrak{sl}_n(K)$. Rank-one square-zero matrices also span $\mathfrak{sl}_n(K)$.*

Proof. For the first assertion, diagonal conjugation has one-dimensional weight spaces KE_{ij}, $i \neq j$, and its weight-zero space is diagonal. If an invariant subspace contains a matrix with a nonzero off-diagonal entry, extraction of its character component gives E_{ij} in that subspace. This extraction is elementary: distinct Laurent monomials on $(K^{\times})^n$ are linearly independent, and evaluation at finitely many diagonal matrices recovers their coefficient vectors by linear algebra. Permutation conjugation gives every off-diagonal matrix unit. Conjugating E_{ij} by $I_n + tE_{ji}$ gives

$$E_{ij} + t(E_{jj} - E_{ii}) - t^2 E_{ji},$$

so the diagonal differences belong as well. If the initial nonzero matrix is diagonal, it is nonscalar because its trace is zero; conjugation by $I_n + tE_{ij}$ for unequal diagonal entries yields a nonzero off-diagonal component. This proves the first assertion.

The span of the conjugates of a nonscalar matrix is invariant. The linear span of differences of conjugates is a nonzero invariant subspace of $\mathfrak{sl}_n(K)$, since a matrix fixed by every similarity is scalar. Applying the first assertion proves the next two claims. Finally let e_i be the ith standard column and e_i^* its coordinate functional. The matrix E_{ij} is rank one and square zero for $i \neq j$, and

$$E_{ii} - E_{jj} = (e_i + e_j)(e_i^* - e_j^*) + E_{ij} - E_{ji}.$$

The outer product in this expression is rank one and square zero, because $(e_i^* - e_j^*)(e_i + e_j) = 0$. These matrices span $\mathfrak{sl}_n(K)$. □

Lemma 2.2. *If an ordinary multilinear polynomial has a value of nonzero trace on $M_n(K)$, its image contains every scalar matrix.*

Proof. Choose a primitive nth root of unity $\zeta \in K$. Define D to be diagonal with entries $1, \zeta, \ldots, \zeta^{n-1}$, and define the cyclic shift R by $Re_i = e_{i+1}$ with cyclic indices. They satisfy $DR = \zeta RD$, including at the cyclic boundary, and therefore

$$(D^a R^b)(D^c R^d) = \zeta^{-bc} D^{a+c} R^{b+d}.$$

Exponents are taken modulo n. For each fixed b, these matrices occupy one cyclic diagonal and the coefficient matrix in a is the invertible Vandermonde matrix $(\zeta^{ai})_{a,i=0}^{n-1}$. Different b give disjoint diagonals. Hence the n^2 matrices form a basis. Their trace is zero for $b \neq 0$; when $b = 0$ it is $\sum_{i=0}^{n-1} \zeta^{ai}$, equal to zero unless $a = 0$, in which case it is $n \neq 0$. Expand the inputs of a value of nonzero trace in this basis. Some basis-input evaluation has nonzero trace. All its words have the same pair of total indices, so that evaluation is a nonzero scalar matrix. Scaling one input gives the assertion. □

Lemma 2.3. *Every hyperplane in $M_n(K)$ distinct from $\mathfrak{sl}_n(K)$ meets every nonscalar conjugacy class.*

Proof. Write the hyperplane as $\{X : \operatorname{tr}(AX) = 0\}$ with A nonscalar. For a nonscalar T, some conjugate C of A does not commute with T: Lemma 2.1 shows that the span of conjugates of A contains $\mathfrak{sl}_n(K)$, whose common centralizer consists of scalars. By Lemma 2.1, rank-one square-zero matrices span $\mathfrak{sl}_n(K)$, so there is such a matrix N with $\operatorname{tr}([C,T]N) \neq 0$. The polynomial

$$t \longmapsto \operatorname{tr}\big(T(I_n + tN)C(I_n - tN)\big)$$

has nonzero derivative at zero. It has a root in K, and $I_n - tN$ is the inverse of $I_n + tN$. Thus a conjugate of C is orthogonal to T, which is equivalent to a conjugate of T lying in the given hyperplane. □

We shall also need a codimension-two version. We use the following consequence of de Bondt's block theorem [14, Theorem 3.3 and Corollary 3.2]. If $S \subset M_n(K)$, $0 < \dim S < n$, and $I_n \notin S$, put

$$r + 1 = \dim_{K(x)} \operatorname{span}_{K(x)} \{Cx : C \in S + KI_n\},$$

Here x is the column of independent indeterminates $x_1, \ldots, x_n$, and $K(x)$ is their rational function field. After a change of basis, projection of S onto the upper-right $r \times (n-r)$ block is injective. In particular $1 \leq r < n$.

Theorem 2.4. *Suppose $n \geq 3$ and I_n, A, B are linearly independent. Every nonscalar conjugacy class O satisfies*

$$\{(\operatorname{tr}(AC), \operatorname{tr}(BC)) : C \in O\} = K^2.$$

Proof. First suppose $n = 3$, so $A, B \in M_3(K)$ and I_3, A, B are independent. Consider a nonscalar semisimple $T \in M_3(K)$ and prescribed values a, b. Choose c with $\tau = \operatorname{tr}(T + cI_3) \neq 0$, and set

$$A' = A - \frac{a + c \operatorname{tr} A}{\tau} I_3, \qquad B' = B - \frac{b + c \operatorname{tr} B}{\tau} I_3.$$

The block theorem applied to $\operatorname{span}\{A', B'\}$ gives $r = 1$ or 2. The spectrum of T admits a disjoint split of sizes $r, 3 - r$. For diagonal blocks D_1, D_2 with these spectra, the lower-left block E in $Z = \left(\begin{smallmatrix} D_1 & 0 \\ E & D_2 \end{smallmatrix}\right)$ can solve $\operatorname{tr}(A'Z) = \operatorname{tr}(B'Z) = 0$: the coefficient map in E is dual to the injective upper-right projection. Indeed, for $S = \operatorname{span}\{A', B'\}$ the map $E \mapsto (T \mapsto \operatorname{tr}(T_{12}E))$ onto S^* is surjective, since $T \mapsto T_{12}$ is injective and the rectangular trace pairing is nondegenerate. It solves the two affine equations after the diagonal contributions are moved to the right side. Solve

$$XD_1 - D_2X = E.$$

The diagonal D_1, D_2 have disjoint spectra, so each entry is solved by division by the corresponding nonzero eigenvalue difference. The matrix $\left(\begin{smallmatrix} I & 0 \\ X & I \end{smallmatrix}\right)$ conjugates $\operatorname{diag}(D_1, D_2)$ to $\left(\begin{smallmatrix} D_1 & 0 \\ E & D_2 \end{smallmatrix}\right)$. Choosing the blocks for $T + cI_3$ gives $Z \sim T + cI_3$. Since $\operatorname{tr} Z = \tau$,

$$\operatorname{tr}(AZ) = \frac{a + c \operatorname{tr} A}{\tau} \operatorname{tr} Z = a + c \operatorname{tr} A,$$

and the same equation holds for B. Thus $Z - cI_3 \sim T$ has pairings a, b.

For arbitrary $n \geq 3$, find a three-dimensional compression on which I_n, A, B remain independent. Put $e = \max_x \dim \operatorname{span}\{x, Ax, Bx\}$. If $e = 3$, take the span of a witnessing triple and any projection onto it. A relation between the three compressed operators, applied to x, is zero only if all coefficients are zero. The case $e = 1$ would make A, B scalar. If $e = 2$, apply the block theorem to $S = \operatorname{span}\{A, B\}$, for which $r = 1$. Its two upper-right rows are independent. Two of their columns detect that independence; compress onto those two coordinate directions and the first coordinate. The compressed identity and the two compressed operators are independent. Denote the compression space by W and its projection by P. In the case $e = 2$, the compression is onto the span of e_1, e_j, e_k, where the 2×2 minor of the two upper-right rows in columns j, k is nonzero. In a relation among the compressed operators, those two off-diagonal entries force the coefficients of A, B to vanish; the diagonal identity entry forces the last coefficient to vanish. Here the compressed operator on W is $A_W = PA|_W$, and likewise for B.

For a nonscalar semisimple T in dimension n, choose three eigenvalues that are not all equal. Put their 3×3 diagonal matrix on W, and put the remaining eigenvalues in a diagonal matrix D on $\ker P$. The compressed constraints are independent modulo the identity. The preceding argument gives a conjugate on W with desired pairings

$$a - \operatorname{tr}(A_{22}D), \qquad b - \operatorname{tr}(B_{22}D),$$

where A_{22}, B_{22} are the complementary blocks. The resulting block diagonal target is conjugate to T and has the required pairings.

The minimal nonzero nilpotent orbit also attains every pair. Seek $C = uv$ with $u \neq 0$, $v \neq 0$, and

$$v(u) = 0, \qquad v(Au) = a, \qquad v(Bu) = b.$$

For $(a, b) \neq (0, 0)$, if $e = 3$, choose a witnessing u and prescribe the three values of v. If $e = 2$, use the independent upper-right rows above: take $v = e_1^*$ and choose $u \in \operatorname{span}\{e_2, \ldots, e_n\}$ solving the last two equations. For $(a, b) = (0, 0)$, take an eigenvector u of A and a nonzero functional annihilating u, Au, Bu. Such a functional exists because this span has dimension

at most two. In each case C is a nonzero rank-one square-zero matrix. Its scalar translate attains every pair as well.

Finally, let O be a nonscalar nonsemisimple orbit, $d = \dim O$, and $X = \overline{O}$. Let F be the affine subspace defined by the two prescribed pairings. If O has several eigenvalues, its nonscalar semisimple part lies in X and meets F. If it has one eigenvalue λ, its closure contains the translate by λI_n of the minimal nonzero nilpotent orbit. For the latter inclusion, take a Jordan representative $\lambda I_n + N$ and choose one nonzero superdiagonal Jordan entry. In each chain assign integer weights to the basis vectors so that consecutive weight differences are positive at every edge except the chosen edge, where the difference is zero. Conjugation by the corresponding diagonal powers of a parameter scales those edges by positive powers and leaves the chosen edge unchanged. Specialization at zero gives a nonzero rank-one nilpotent plus λI_n. The same construction with every edge assigned a positive difference gives the semisimple part in the former case. Thus $X \cap F$ is nonempty.

There are finitely many boundary orbits in $X \setminus O$, since the characteristic polynomial is fixed. Write $Z(T) = \{C : CT = TC\}$ for the vector-space centralizer. If the Jordan blocks for eigenvalue λ have lengths $r_1, r_2, \ldots$, then a commuting block between lengths r_i, r_j has dimension $\min(r_i, r_j)$: it is a homomorphism of the cyclic $K[t]$-modules $K[t]/(t^{r_i})$ and $K[t]/(t^{r_j})$. Blocks for distinct eigenvalues contribute zero. If $c_{\lambda,j}$ is the number of λ-blocks of length at least j, then

$$\dim Z(T) = \sum_{\lambda} \sum_{i,k} \min(r_i, r_k) = \sum_{\lambda,j} c_{\lambda,j}^2, \qquad \dim O_T = n^2 - \dim Z(T).$$

The positive $c_{\lambda,j}$ sum to n, and $c_{\lambda,j}^2 \equiv c_{\lambda,j} \pmod 2$, so this orbit dimension is even. The closure X is irreducible of dimension d, as it is the closure of the image of the irreducible variety $\mathrm{GL}_n(K)$. Each boundary orbit has smaller dimension, hence dimension at most $d-2$. On every nonscalar orbit, $C \mapsto \mathrm{tr}(AC)$ is nonconstant. Indeed, the span of its orbit differences is a nonzero invariant subspace of $\mathfrak{sl}_n(K)$, hence all of $\mathfrak{sl}_n(K)$; a constant pairing would make A scalar. It follows that

$$\dim((X \setminus O) \cap F) \leq \max\{0, d-3\}.$$

Every component of $X \cap F$ has dimension at least $d-2$, by the height theorem. For a nonscalar T, every $c_{\lambda,j}$ is at most $n-1$; the only excluded partition with a column of length n is a scalar matrix. Among positive integers of sum n with this bound, the sum of squares is at most $(n-1)^2 + 1$, since transferring a unit from a smaller to a larger part increases that sum. Thus $d \geq 2n-2 \geq 4$. Therefore $X \cap F$ cannot be contained in the boundary, and $O \cap F \neq \varnothing$. □

Corollary 2.5. *Suppose $n \geq 3$, p has a value of nonzero trace, and one slice of p has rank at least $n^2 - 2$. Then $p(M_n(K)) = M_n(K)$.*

Proof. In the same frozen-input space, intersect the rank open set with the open set on which the slice has nonzero trace. Its image H has codimension at most two and is not contained in $\mathfrak{sl}_n(K)$. Its orthogonal space for the trace pairing excludes I_n. Theorem 2.4, or Lemma 2.3 in codimension one, shows that H meets every nonscalar orbit. Polynomial images are invariant under simultaneous conjugation of the inputs. Lemma 2.2 supplies the scalar values. □

3 Two-cycle slices

Fix all but two inputs, denoted B, C, to diagonal matrices. If there are s frozen inputs, order them by their original labels and write $z_i \in K^s$ for their diagonal entries in row i.

Define $F(u,v,w)$ by summing over words with B before C: multiply the word coefficient by the coordinates of u for diagonal letters before B, of v for letters between B, C, and of w for letters after C. Define G by the same rule for words with C before B. Empty products are one. For $1 \le i, j \le n$, expansion along matrix entries gives

$$p_{ij} = \sum_{h=1}^{n} F(z_i, z_h, z_j) B_{ih} C_{hj} + \sum_{h=1}^{n} G(z_i, z_h, z_j) C_{ih} B_{hj}. \tag{2}$$

For example, with two diagonal inputs A, D, the word $ABDC$ contributes $a_i d_h$ to F, and $ACDB$ contributes $a_i d_h$ to G. For independent tuples $u, v \in K^s$, define

$$P(u,v) = F(u,v,u), \quad Q(u,v) = G(v,u,v), \quad \Delta(u,v) = Q(u,v)Q(v,u) - P(u,v)P(v,u).$$

Lemma 3.1. *Every nonscalar $T \in M_3(K)$ has a zero one-dimensional compression whose complementary two-dimensional compression is diagonalizable.*

Proof. Choose w, ϕ with $\phi(w) = 1$, $\phi(Tw) = 0$, so the decomposition $Kw \oplus \ker \phi$ has a zero first diagonal entry. Write $T = \left(\begin{smallmatrix} 0 & r \\ c & S \end{smallmatrix}\right)$ in this decomposition. Then $\phi(T^2 w) = rc$, $\operatorname{tr}(T^2) = 2rc + \operatorname{tr}(S^2)$, and $\operatorname{tr} S = \operatorname{tr} T$. Defining $e_2(T)$ by

$$\det(tI_3 - T) = t^3 - (\operatorname{tr} T)t^2 + e_2(T)t - \det T,$$

we have $e_2(T) = ((\operatorname{tr} T)^2 - \operatorname{tr}(T^2))/2$ and $\det S = ((\operatorname{tr} S)^2 - \operatorname{tr}(S^2))/2$. Consequently

$$\operatorname{tr} S = \operatorname{tr} T, \qquad \det S = e_2(T) + \phi(T^2 w).$$

When the minimal polynomial has degree three, choose w cyclic. The last displayed functional value can then be prescribed freely, giving distinct eigenvalues for S. If the minimal polynomial is $(t-\lambda)(t-\mu)$ with $\lambda \neq \mu$, label the eigenvalues so that their algebraic multiplicities are two and one respectively. Then $T^2 = (\lambda+\mu)T - \lambda\mu I_3$ gives discriminant μ^2 for S. For $\mu \neq 0$, choose any w, Tw independent and the indicated ϕ. For $\mu = 0$, use its eigenspace as Kw and the λ-eigenspace as complement. If $T = \lambda I_3 + N$, $N^2 = 0 \neq N$, the same calculation gives discriminant λ^2. For $\lambda = 0$, take $Kw = \operatorname{im} N$; projection along w makes the complementary compression zero. These exhaust the nonscalar cases. □

Lemma 3.2. *Fix a pairing of all coordinates in even dimension $n \ge 4$. The subspace defined by zero entries in the two ordered positions joining each pair meets every conjugacy class. In odd dimension $n \ge 3$, fix a pairing leaving one coordinate unpaired and also set its diagonal entry to zero. This subspace meets every nonscalar conjugacy class.*

Proof. For $n = 2r \ge 4$, put the target in Jordan form and pair positions $i, i+r$. Paired positions are nonadjacent, and both joining entries vanish. For odd $n \ge 5$, choose w, Tw independent and $\phi(w) = 1$, $\phi(Tw) = 0$. The complementary compression has even dimension at least four. Put it in Jordan form and pair its nonadjacent positions. All entries joining Kw to its complement are unrestricted. Any two coordinate pairings are related by a permutation, so this also proves the assertion for a fixed prescribed pairing. For $n = 3$ use Lemma 3.1 and pair the two eigenvectors of the complementary compression. □

Proposition 3.3. *If Δ is not the zero polynomial in (2), then $p(M_n(K)) = M_n(K)$ for every $n \ge 3$.*

Proof. First suppose F, G are both nonzero. Let $B_{i,\sigma(i)} = b_i \neq 0$, where σ consists of two-cycles and, for odd n, one singleton. Set $Z = BC$. Direct inversion gives

$$p_{ij} = F(z_i, z_{\sigma(i)}, z_j) Z_{ij} + \frac{b_{\sigma^{-1}(j)}}{b_{\sigma^{-1}(i)}} G(z_i, z_{\sigma^{-1}(j)}, z_j) Z_{\sigma^{-1}(i),\sigma^{-1}(j)}. \tag{3}$$

Thus the operator preserves orbits of the simultaneous permutation of the two indices. The diagonal orbit of each two-cycle has determinant Δ up to a nonzero factor. Set the source entries in the other internal orbit equal to zero. This makes the two joining output entries zero.

Give all weights in each cycle a common value, with independent nonzero values for different cycles. On an entry orbit of length d between cycles of weights w_1, w_2, the coefficient matrix is diagonal plus (w_2/w_1) times a weighted cyclic permutation. Its determinant is

$$\prod_{(i,j)} F(z_i, z_{\sigma(i)}, z_j) + (-1)^{d-1}(w_2/w_1)^d \prod_{(i,j)} G(z_i, z_{\sigma^{-1}(j)}, z_j),$$

where each product runs over that entry orbit. Indeed, in the determinant expansion a row can select its diagonal or cyclic entry; selecting one cyclic entry forces that choice around the entire orbit. Between two two-cycles, every factor is a nonzero polynomial on three independent vertex tuples. For a two-cycle with tuples u, v and a singleton with tuple w, the first product in the cycle-to-singleton direction is $F(u, v, w)F(v, u, w) \neq 0$; in the reverse direction the second is $G(w, u, v)G(w, v, u) \neq 0$. The distinct weight powers therefore give a nonzero determinant in every case. Their finite product is nonzero, so all needed determinants can be chosen nonzero simultaneously. Set the singleton source diagonal to zero. The slice then contains the pattern of Lemma 3.2, proving coverage of nonscalar targets. The invertible two-cycle diagonal map gives a value of nonzero trace already in size two. Corner embedding and Lemma 2.2 give every scalar target in all sizes.

If $F = 0$, the condition on Δ implies $Q \neq 0$. Choose $\sigma(i) = i - 1$ modulo n. The map is now a weighted permutation of entries. For a diagonal output its coefficient is $G(u, v, u)$, a nonzero polynomial. For a first-superdiagonal output, its three vertex indices are distinct because $n \geq 3$, so its coefficient is also nonzero. Choose all needed coefficients nonzero, solve for any upper bidiagonal target, and set every unused source entry to zero. Every matrix has a Jordan representative of this form. If $G = 0$, the same reverse-cycle construction uses the F term, whose intermediate vertex is again distinct at the first-superdiagonal positions. □

4 Identity reductions and scalar shifts

Write $\mathrm{ad}_A(Y) = [A, Y]$ for the commutator operator. The notation $p(X_i = I)$ means the free polynomial obtained by replacing the named generator by the unit; in matrix evaluation this unit is I_n. Call a multilinear polynomial *proper* if every such substitution is zero. For a cubic, write the coefficient vector in the order $ABC, ACB, BAC, BCA, CAB, CBA$ as (a, b, c, d, e, f). The three identity substitutions give

$$a + c + d = b + e + f = a + b + c = d + e + f = a + b + e = c + d + f = 0.$$

Solving gives $(a, b, c, d, e, f) = (a, -a - c, c, -a - c, c, a)$. These are exactly the coefficients of $a[A, [B, C]] + c[B, [A, C]]$. Thus the proper subspace is

$$\mathrm{span}\{[A, [B, Y]], [B, [A, Y]]\}. \tag{4}$$

We use the Lie algebra $\mathfrak{sl}_2(K)$ with generators $h = \mathrm{diag}(1, -1)$, $e = E_{12}$, $f = E_{21}$. Its n-dimensional irreducible representation ρ_n can be realized on a basis $w_0, \ldots, w_{n-1}$ by

$$\rho_n(h)w_i = (n - 1 - 2i)w_i, \quad \rho_n(e)w_i = i(n - i)w_{i-1}, \quad \rho_n(f)w_i = w_{i+1},$$

with vectors outside the stated range equal to zero. These formulas satisfy the three commutator relations of h, e, f.

Lemma 4.1. *Let $Q = (q_{ab})$ be a 3×3 matrix over K with rank at least two and nonzero symmetric part $(Q + Q^{\mathsf{T}})/2$. For each $n \geq 2$, there are $x_1, x_2, x_3 \in \mathfrak{sl}_2(K)$ such that, in the irreducible n-dimensional representation ρ_n of $\mathfrak{sl}_2(K)$,*

$$M = \sum_{a,b=1}^{3} q_{ab} \operatorname{ad}_{\rho_n(x_a)} \operatorname{ad}_{\rho_n(x_b)}$$

is invertible on $\mathfrak{sl}_n(K)$.

Proof. View Q as the bilinear form $Q(u, v) = u^{\mathsf{T}} Q v$ on K^3. It has a nondegenerate two-dimensional restriction with nonzero symmetric part. To prove the first assertion, suppose all two-dimensional restrictions were degenerate. Choose u with $Q(u, u) = 1$ and put $V = \{v : Q(u, v) = 0\}$. Degeneracy on $\operatorname{span}\{u, v\}$ implies $Q(v, v) = 0$ for $v \in V$. Thus $Q|_V$ is alternating. If $Q(v, w) = b \neq 0$, the determinant on $\operatorname{span}\{u + tv, w\}$ is $-tb(Q(w, u) - tb)$, a contradiction. Hence $Q|_V = 0$, forcing $\operatorname{rank} Q = 1$. Nondegeneracy and nonzero symmetric part are nonempty open conditions on pairs of vectors, so they hold together.

For $x, y \in \mathfrak{sl}_2(K)$ and basis columns u, v of such a plane, substituting $x_a = u_a x + v_a y$ replaces the coefficient matrix by the restriction of Q to that plane. A change of basis in the plane is therefore an allowed substitution. By congruence, the restricted matrix has one of the forms

$$\begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}, \quad \alpha\beta(\alpha + \beta) \neq 0; \qquad \begin{pmatrix} s & k \\ -k & 0 \end{pmatrix}, \quad sk \neq 0.$$

These follow by choosing an isotropic basis when the symmetric part has rank two, and by diagonalizing it when it has rank one. It suffices to treat the corresponding operators $\alpha AB + \beta BA$ and $sA^2 + k[A, B]$ in every nontrivial integer-spin representation.

Let V_j be the irreducible module of dimension $2j + 1$, $j \in \mathbb{Z}_{\geq 1}$. Denote the actions of h, e, f by H, E, F, and regard A, B in the preceding operator expressions as actions of elements of $\mathfrak{sl}_2(K)$. Choose a basis v_m, with integer $-j \leq m \leq j$, with

$$Hv_m = 2mv_m, \quad Ev_m = (j - m)v_{m+1}, \quad Fv_m = (j + m)v_{m-1},$$

and with vectors beyond the endpoints zero. In the first case take $A = H$, $B = H + t(E + F)$. At $t = 0$ the operator is $(\alpha + \beta)H^2$, with its only kernel on weight zero. The scalar Schur complement on that weight has coefficient

$$-\frac{2\alpha\beta j(j + 1)}{\alpha + \beta} t^2$$

at order two. The first-order map sends v_0 to $2\alpha j(v_1 - v_{-1})$; its coefficient from $v_{\pm 1}$ back to v_0 is $\pm 2\beta(j + 1)$. The inverse of the nonzero-weight block at zero has coefficient $1/(4(\alpha + \beta))$ at both v_1, v_{-1}. The negative of the product of these three maps gives the displayed coefficient. In the second case take $A = H$, $B = t(E + F)$; the operator is $sH^2 + 2kt(E - F)$, and the corresponding coefficient is $2k^2 j(j + 1)t^2/s$. Here the outgoing coefficients are $2kj, -2kj$, the incoming coefficients are $-2k(j + 1), 2k(j + 1)$, and the inverse diagonal coefficients are $1/(4s)$. Both coefficients are nonzero, so each determinant is a nonzero polynomial in the original three x_a.

Under the action $x \cdot Y = [\rho_n(x), Y]$, the nonzero matrices $\rho_n(e)^j$, $0 \leq j < n$, are highest-weight vectors of weights $2j$. The standard highest-weight calculation

$$EF^r v = r(2j - r + 1)F^{r-1} v$$

for $Ev = 0$, $Hv = 2jv$, shows that each generates an irreducible module of dimension $2j + 1$. $F^{2j+1} v = 0$: it would otherwise be a highest-weight vector of weight $-2j - 2$,

and the same identity would produce an infinite nonzero descending sequence, impossible in finite dimension. For $0 \le r \le 2j$ the displayed nonzero coefficients show that $F^r v \neq 0$. The weights are distinct, and any nonzero invariant subspace contains a weight vector; applying E repeatedly recovers v. This proves irreducibility. Distinct highest weights make these simple modules nonisomorphic, hence their sum is direct; their dimensions add to $\sum_{j=0}^{n-1}(2j+1) = n^2$. Consequently

$$M_n(K) = V_0 \oplus V_1 \oplus \cdots \oplus V_{n-1}.$$

Here $V_0 = KI_n$. On any nontrivial simple summand, trace is an invariant linear functional and hence zero; their sum is therefore $\mathfrak{sl}_n(K)$. Intersect the finitely many nonvanishing determinant conditions to obtain the required tuple. □

Lemma 4.2. *Suppose a quartic p has zero double-identity reductions and proper single-identity reductions. Fix a free input Y, and denote the other inputs, in their original order, by x_1, x_2, x_3. Suppose scalars $\lambda_1, \lambda_2, \lambda_3 \in K$ make the coefficient of t after substituting $x_a + t\lambda_a I$ equal to*

$$\sum_{a,b=1}^{3} q_{ab} \operatorname{ad}_{x_a} \operatorname{ad}_{x_b}(Y),$$

where the scalar matrix $Q = (q_{ab})$ has rank at least two and nonzero symmetric part. If p has nonzero formal trace, then $p(M_n(K)) = M_n(K)$ for every $n \ge 3$. The same conclusion holds if an identity reduction is

$$\alpha[A,[B,Y]] + \beta[B,[A,Y]], \qquad \alpha\beta(\alpha+\beta) \neq 0.$$

Proof. After the shifts the slice is exactly $L_0 + tM$: the term indexed by a set of two or more shifts contains a double-identity reduction and is zero. Choose $x_a' \in \mathfrak{sl}_2(K)$ using Lemma 4.1 and freeze $x_a = \rho_n(x_a')$, so M is invertible on $\mathfrak{sl}_n(K)$. Restrict the source to $\mathfrak{sl}_n(K)$ and project the target by $T \mapsto T - (\operatorname{tr} T/n)I_n$. The resulting $(n^2-1) \times (n^2-1)$ matrix has determinant with leading coefficient $\det(M|_{\mathfrak{sl}_n(K)}) \neq 0$. For some t it is invertible, and the full slice has rank at least $n^2 - 1$. The same rank condition is therefore nonempty open in the full frozen-input space. The trace-nonzero slice condition is nonempty open there as well. Their intersection gives a full image or a hyperplane distinct from $\mathfrak{sl}_n(K)$. Apply Lemmas 2.3 and 2.2. For the final assertion, the cubic itself gives a rank-n^2-1 slice by Lemma 4.1, using the matrix with the two off-diagonal entries α, β. □

Lemma 4.3. *Suppose a quartic has identity reductions*

$$p(A,B,C,I) = \lambda[A,[B,C]], \qquad p(I,B,C,D) = \mu[D,[B,C]], \quad \lambda \neq 0,$$

and its restriction with A, D diagonal has nonzero formal trace. Then it is surjective on every $M_n(K)$, $n \ge 3$.

Proof. For diagonal A, D, the coefficient of $B_{ij}C_{ji}$ in the ith output is $\alpha(a_i - a_j)(d_i - d_j)$, and that of $C_{ij}B_{ji}$ is $\beta(a_i - a_j)(d_i - d_j)$. Indeed, the first reduction makes each coefficient vanish at $d_i = d_j$, since a commutator with diagonal A has zero diagonal. The second gives vanishing at $a_i = a_j$. Each coefficient is bilinear in the a and d entries, so it is a scalar multiple of $(a_i - a_j)(d_i - d_j)$. Nonzero trace means $\alpha + \beta \neq 0$. If $\alpha^2 \neq \beta^2$, Proposition 3.3 applies. Otherwise $\alpha = \beta \neq 0$.

Take $B = \sum_{i=1}^{n-1} E_{i,i+1}$ and $C = \sum_{i=1}^{n-1} 3^i E_{i+1,i}$. With $c_i = 3^i$ for $1 \le i < n$ and $c_0 = c_n = 0$, the diagonal entries of $[B,C]$ are $c_i - c_{i-1}$ and are pairwise distinct. Choose D diagonal with $d_i \neq d_{i+1}$. On diagonal A, edge i contributes $\alpha c_i(d_i - d_{i+1})(a_i - a_{i+1})$ to both

endpoint coordinates. The resulting map is the product of the surjective difference map, an invertible diagonal map, and the injective signless incidence map of a path. Its rank is $n-1$, with left kernel $(1,-1,1,-1,\dots)$, so it has nonzero trace.

Replacing D by $D+tI_n$ adds $t\lambda[A,[B,C]]$. The band of offset r, $-(n-1)\le r\le n-1$, is $\operatorname{span}\{E_{i,i+r} : 1\le i, i+r\le n\}$. Multiplication adds offsets, and B, C, D have offsets $1,-1,0$, so each band is preserved; on every nonzero band this leading map is invertible, since the diagonal entries of $[B,C]$ are distinct. For generic t the full slice has rank n^2-1 and nonzero trace. Lemmas 2.3 and 2.2 finish the proof. □

5 Proper quartics

Proposition 5.1. *Every proper multilinear quartic with nonzero formal trace is surjective on $M_n(K)$ for every $n\ge 3$.*

Proof. Let $\operatorname{sgn}(\sigma)$ be 1 for an even permutation and -1 for an odd permutation. For $(X_1,X_2,X_3,X_4)=(A,B,C,D)$, define

$$s_4(A,B,C,D)=\sum_{\sigma\in S_4}\operatorname{sgn}(\sigma)X_{\sigma(1)}X_{\sigma(2)}X_{\sigma(3)}X_{\sigma(4)}.$$

Cyclic rotation of four distinct letters changes the sign and preserves matrix trace, so $\operatorname{tr} s_4 = 0$. Some identification of two inputs of p has nonzero trace. Indeed, if the four-linear trace form vanished whenever two inputs were equal, replacing that common input by a sum would show that exchanging those two inputs negates the form. It would therefore be alternating. Alternating the words of p, and dividing by 4!, gives a scalar multiple of s_4 with the same trace form. This contradicts the trace nonvanishing of p.

A proper polynomial of multidegree $(2,1,1)$ in A,B,C has the form

$$g=\alpha[A,B][A,C]+\beta[A,C][A,B]+\gamma[A,[A,[B,C]]]+\delta[[A,[A,B]],C]. \tag{5}$$

For a word, its central-translation derivative in a letter is the sum of the words obtained by deleting one occurrence of that letter. This is the coefficient of t after replacing that input by itself plus tI_n. A repeated specialization of a proper multilinear polynomial has all three such derivatives zero. The coefficient matrix of these maps on the 12 words has rank eight. The following explicit coefficient columns and nonzero minors verify that its kernel is exactly the span in (5).

word	α	β	γ	δ
$AABC$	0	0	1	1
$AACB$	0	0	-1	0
$ABAC$	1	0	0	-2
$ABCA$	-1	0	-2	0
$ACAB$	0	1	0	0
$ACBA$	0	-1	2	0
$BAAC$	-1	0	0	1
$BACA$	1	0	0	0
$BCAA$	0	0	1	0
$CAAB$	0	-1	0	-1
$CABA$	0	1	0	2
$CBAA$	0	0	-1	-1

For the derivative matrix, order the deleted letter as A, B, C and each target word lexicographically, numbering rows and columns from zero. Its minor on rows $0,1,2,3,4,5,6,9$ and columns $0,1,2,3,4,5,6,8$ is -4. The four printed columns have minor 1 on rows $0,1,2,4$, and the derivative matrix annihilates them. Hence its kernel has exactly this basis. The trace of (5) is $(\alpha+\beta)\,\mathrm{tr}([A,B][A,C])$, because the last two terms are commutators and trace is cyclic. Its nonzero value forces $\alpha+\beta \neq 0$.

For a permutation σ of $\{1,\dots,n\}$, take $A = \mathrm{diag}(a_i)$ and the weighted permutation matrix B with $B_{i,\sigma(i)} = b_i \neq 0$ and every other entry zero. Set $Z = BC$. The map in C is

$$g_{ij} = F_{ij}Z_{ij} - H_{ij}\frac{b_{\sigma^{-1}(j)}}{b_{\sigma^{-1}(i)}}Z_{\sigma^{-1}(i),\sigma^{-1}(j)},$$

where

$$\begin{aligned} F_{ij} &= \alpha(a_i - a_{\sigma(i)})(a_{\sigma(i)} - a_j) + \gamma(a_i - a_j)^2 + \delta(a_i - a_{\sigma(i)})^2, \\ H_{ij} &= \gamma(a_i - a_j)^2 + \delta(a_{\sigma^{-1}(j)} - a_j)^2 - \beta(a_i - a_{\sigma^{-1}(j)})(a_{\sigma^{-1}(j)} - a_j). \end{aligned}$$

The two-cycle diagonal determinant is nonzero unless $(\delta-\alpha)^2 = (\beta+\delta)^2$. Equality together with $\alpha+\beta \neq 0$ forces $2\delta = \alpha - \beta$. Outside this exceptional case, the two-cycle proof above applies with the single diagonal input A; scalar targets follow from the original multilinear polynomial.

In the exceptional case retain the original coefficients, with $2\delta = \alpha-\beta$ and $\alpha+\beta \neq 0$. For a reverse cycle of odd length $\ell \geq 3$, put $x_i = a_i - a_{i+1}$ with cyclic indices. These differences range over the hyperplane $\sum_i x_i = 0$. The diagonal determinant is

$$\left((\delta-\alpha)^\ell - (\beta+\delta)^\ell\right)\prod_i x_i^2 = -2\left(\frac{\alpha+\beta}{2}\right)^\ell \prod_i x_i^2,$$

a nonzero polynomial. Define binary quadratic polynomials by

$$f(u,v) = \gamma u^2 - \alpha uv + (\delta-\alpha)v^2, \qquad h(u,v) = (\beta+\delta)u^2 + \beta uv + \gamma v^2.$$

Both are nonzero, since $\delta - \alpha = -(\alpha+\beta)/2$ and $\beta+\delta = (\alpha+\beta)/2$. For a reverse cycle $\sigma(i) = i-1$, the coefficients on the cyclic first-superdiagonal are

$$F_{i,i+1} = f(x_i, x_{i-1}), \qquad H_{i,i+1} = h(x_{i+1}, x_i).$$

For example, $a_i - a_{i-1} = -x_{i-1}$ and $a_{i-1} - a_{i+1} = x_{i-1} + x_i$ in the formula for $F_{i,i+1}$. Since the product of the F coefficients can be reindexed cyclically, the first-superdiagonal determinant is

$$\prod_i f(x_{i+1}, x_i) - \prod_i h(x_{i+1}, x_i), \qquad \sum_i x_i = 0. \tag{6}$$

For $\ell \geq 5$, work in $K[x_1,\dots,x_\ell]/(\sum_i x_i)$, a polynomial ring in $\ell-1$ variables and hence a unique factorization domain. Factor the binary quadratics into linear factors over K. Two factors supported on at most two cyclic coordinates can be proportional in this quotient only when they were already proportional before quotienting: their difference has support on at most four coordinates. Factors involving both endpoints of an edge must therefore match on that edge; only individual-coordinate factors can pass between edges. Cancel the matching two-endpoint factors. The remaining factors are powers of the two individual coordinates; equality of the two degrees forces opposite exponents in their quotient. An identity in (6) would thus imply, for $c \in K^\times$ and an integer k,

$$f(u,v)/h(u,v) = c(u/v)^k, \quad -2 \leq k \leq 2, \quad c^\ell = 1.$$

The coefficient comparisons can be listed explicitly. For $k = 1$, the identity $vf = cuh$ forces $\beta + \delta = 0$ and $\delta - \alpha = 0$, contrary to $\alpha + \beta \neq 0$. For $k = 2$, comparison of v^4 in $v^2 f = cu^2 h$ again gives $\delta - \alpha = 0$, a contradiction. For $k = -2$, comparison of $u^4, u^3 v, uv^3$ in $u^2 f = cv^2 h$ gives $\gamma = \alpha = \beta = 0$, also a contradiction. For $k = 0$, $f = ch$ gives $-\alpha = c\beta$, $\gamma = c(\beta + \delta)$, $\delta - \alpha = c\gamma$. Thus

$$\frac{\delta - \alpha}{\beta + \delta} = c^2 = -1,$$

and $c^\ell \neq 1$ for odd ℓ. For $k = -1$, $uf = cvh$ gives

$$\gamma = 0, \quad -\alpha = c(\beta + \delta), \quad \delta - \alpha = c\beta, \quad \delta(1 + c) = 0.$$

If $\delta = 0$, the quotient $(\delta - \alpha)/(\beta + \delta)$ equals $c = -1$; otherwise $c = -1$. Again $c^\ell \neq 1$. This proves nonvanishing for every odd $\ell \geq 5$.

For $\ell = 3$ let $\Phi(u, v)$ be (6) at $(x_1, x_2, x_3) = (u, v, -u - v)$. The coefficient of u^6 is

$$-(\alpha + \beta)\gamma(\delta + \gamma).$$

Here and below a coefficient is taken in the ordinary commutative polynomial ring $K[u, v]$, with the other symbols fixed scalar parameters. An identity $\Phi = 0$ forces $\gamma = 0$ or $\gamma = -\delta$. In both cases, using $2\delta = \alpha - \beta$, its $u^5 v$ coefficient is

$$-\frac{(\alpha - \beta)(\alpha + \beta)^2}{4}.$$

It forces $\alpha = \beta$. Then $\delta = \gamma = 0$ and the $u^4 v^2$ coefficient is $2\alpha^3 \neq 0$. Thus Φ is not the zero polynomial.

Every odd $n \geq 3$ can now use one reverse cycle. Its diagonal and first-superdiagonal maps are invertible for a simultaneous choice of eigenvalues, so the slice contains every upper bidiagonal matrix. For even $n \geq 6$, use two reverse odd cycles of lengths $3, n - 3$. For $n = 4$, use a three-cycle and a singleton. On cross-cycle index orbits of length d, give the cycles independent common weights w_1, w_2. The determinant is

$$\prod_{(i,j)} F_{ij} - (w_2/w_1)^d \prod_{(i,j)} H_{ij},$$

by the same cyclic determinant expansion as in Proposition 3.3. Between two nontrivial cycles the three eigenvalue arguments in each factor are independent. For a three-cycle and singleton, use the F product in the cycle-to-singleton direction and the H product in the reverse direction; those triples are independent as well. In the exceptional case F, H are both nonzero: evaluating their diagonal coefficients gives respectively $\delta - \alpha = -(\alpha + \beta)/2$ and $\beta + \delta = (\alpha + \beta)/2$. Thus cross rectangles are simultaneously invertible.

For $n \geq 6$, put the target in Jordan form with the cycles on consecutive intervals; solve the diagonal and first-superdiagonal bands inside the cycles and all cross rectangles. For $n = 4$, a nonscalar target has a zero one-dimensional compression; put the complementary three-dimensional compression in Jordan form and solve the same bands and cross rectangles, with the singleton source diagonal zero. Nonzero-trace evaluations of (5) in size three embed in every larger size, and Lemma 2.2 gives scalar targets for the original multilinear polynomial. This proves the proposition. □

6 Block transfer from sizes two and three

Lemma 6.1. *Fix a free input position of an ordinary multilinear polynomial p. If p is not an identity on $M_{\min(k,l)}(K)$, then, for generic independent frozen blocks of sizes k, l, the operator on the $k \times l$ rectangle of the free input is invertible.*

Proof. Let m be the degree of p and let r be the free input label. Freeze the other inputs at $\operatorname{diag}(A_i, B_i)$, $A_i \in M_k(K)$, $B_i \in M_l(K)$. On a source rectangle $Y \in M_{k\times l}(K)$ the slice acts as

$$L_{kl}(Y) = \sum_w a_w\, A_{i_1} \cdots A_{i_s}\, Y\, B_{j_1} \cdots B_{j_{m-1-s}},$$

where the word w has the ordered labels $i_1, \ldots, i_s$ before r and $j_1, \ldots, j_{m-1-s}$ after it, and a_w is its coefficient in p. An empty left or right product is I_k or I_l, respectively. Group these operators according to the set S of labels preceding r, and denote their sum for that set by C_S. Some C_S is nonzero: a nonzero evaluation in size $\min(k, l)$ embeds in the rectangle, and its sum is the original slice. Expand the left inputs indexed by S and the right inputs indexed by its complement in separate clock-and-shift bases. Some such homogeneous specialization is nonzero. Every left product is a scalar multiple of one invertible matrix P, and every right product is a scalar multiple of one invertible matrix Q. Hence this specialization is $C_S(X) = cPXQ$, $c \neq 0$, an invertible operator. Scale the left input for each label in S by t, and the right input for each other label by t. The term with prefix set T has degree $m - 1 - |S \triangle T|$, where $\triangle$ denotes symmetric difference. The unique highest coefficient is $t^{m-1}C_S$. The leading coefficient of $\det L_{kl}$ is $\det(C_S) = c^{kl}(\det P)^l(\det Q)^k \neq 0$. This proves the required polynomial nonvanishing. □

Lemma 6.2. *Fix a free input position. Suppose p is not trace zero on $M_2(K)$, and the projection of some slice onto $M_2(K)/KI_2$ has rank at least two. Then a nonempty open set of frozen tuples has slice images meeting every nonscalar 2×2 orbit of nonzero trace.*

Proof. Let A denote the frozen inputs, let F be the rational function field in their independent entries, and let $W \subset M_2(F)$ be the image of the generic slice. The two hypotheses give $\operatorname{tr}|_W \neq 0$ and $\dim(W + FI_2)/FI_2 \geq 2$. Its trace-one section has nonconstant determinant. Indeed, if determinant were constant there, homogeneity would give $\det w = c(\operatorname{tr} w)^2$ on W. For two independent generic free inputs Y, Z, set

$$T = \operatorname{tr}(p(A, Z))p(A, Y) - \operatorname{tr}(p(A, Y))p(A, Z),$$

where A denotes all frozen generic matrices. Then $\operatorname{tr} T = \det T = 0$, so $T^2 = 0$. The algebra of generic matrices together with their characteristic coefficients embeds in its central division algebra of fractions [9, Proposition 1.31]. Both values and both traces defining T belong to that same division algebra. Thus $T = 0$, which would make a generic slice have rank at most one, a contradiction.

Choose $U, V \in W$ with $\operatorname{tr} U \neq 0$, with U, V, I_2 independent, and with determinant nonconstant on the line $U/\operatorname{tr} U + zB$, where $B = \operatorname{tr}(U)V - \operatorname{tr}(V)U$. To justify the choice, nonconstancy on the affine trace-one section gives a line on which determinant varies: a polynomial constant on every line through every point would be constant. Its base point can serve as $U/\operatorname{tr} U$, and its direction is a trace-zero element of W of the displayed form B. Thus determinant variation is a nonempty open condition on pairs U, V. Independence of U, V, I_2 is another nonempty open condition, since the image of W modulo scalars has dimension at least two; so is $\operatorname{tr} U \neq 0$. The pair space is irreducible, and these conditions hold simultaneously. Nonconstancy is the condition that at least one of $\det B$, $\operatorname{tr}(UB)$ is nonzero, by the identity

$$\det\left(\frac{U}{\operatorname{tr} U} + zB\right) = \frac{\det U}{(\operatorname{tr} U)^2} - \frac{z\operatorname{tr}(UB)}{\operatorname{tr} U} + z^2 \det B.$$

The plane $\operatorname{span}\{U, V\}$ excludes I_2, and its trace-one line attains every determinant over K. After scaling, it meets every nonscalar orbit of nonzero trace: a nonscalar 2×2 matrix is

determined up to similarity by its trace and determinant. Write $U = L_A(Y_1)$, $V = L_A(Y_2)$, where L_A is the slice and Y_1, Y_2 are two freely varying source matrices. Independence is witnessed by a 3×3 minor of the three vectorized matrices U, V, I_2. Variation is witnessed by one nonzero coefficient in the determinant formula. Together with $\operatorname{tr} U \neq 0$, their product is a nonzero polynomial in A, Y_1, Y_2, after clearing powers of $\operatorname{tr} U$. Choose a K-point where it is nonzero and fix its Y_1, Y_2. Such a point exists over the infinite field K. The product remains a nonzero polynomial in A, and its nonvanishing defines the asserted nonempty open set. □

Lemma 6.3. *If $T \in M_n(K)$ is nonscalar and $\sum_i \gamma_i = \operatorname{tr} T$, some conjugate of T has diagonal $\gamma_1, \ldots, \gamma_n$.*

Proof. This is Fillmore's theorem [15]; the following construction specifies the similarities. If T is diagonal, choose unequal diagonal entries in positions r, s. Conjugation by $I_n + tE_{rs}$ with $t = (T_{ss} - T_{rr})^{-1}$ produces entry 1 in position (r, s). Thus assume $T_{rs} \neq 0$ for $r \neq s$. Let B be the identity except in row s, where

$$B_{sr} = 0, \quad B_{ss} = T_{rs}, \quad B_{sk} = T_{rk} - 1 \quad (k \neq r, s).$$

Its determinant is T_{rs}. Direct multiplication shows that row r of BTB^{-1} has all off-diagonal entries equal to one: its sth entry is T_{rs}/B_{ss}, and its kth entry for $k \neq r, s$ is $T_{rk} - T_{rs}B_{sk}/B_{ss}$. Replace T by this conjugate. Put $u_r = 0$ and $u_k = \gamma_k - T_{kk}$ for $k \neq r$, and use $C = I_n + ue_r^*$, $C^{-1} = I_n - ue_r^*$. For $k \neq r$,

$$(CTC^{-1})_{kk} = T_{kk} + u_k T_{rk} = \gamma_k.$$

The remaining diagonal entry is $T_{rr} - \sum_{k \neq r} u_k = \gamma_r$ by the trace condition. □

Proposition 6.4. *Suppose a quartic p has nonzero trace on $M_2(K)$. In the same free position suppose that some $M_2(K)$ slice has rank at least two modulo scalars and some $M_3(K)$ slice has rank at least seven. Then $p(M_n(K)) = M_n(K)$ for every $n \geq 3$.*

Proof. In size three, intersect the rank-seven and trace-nonzero opens. Theorem 2.4 makes each resulting slice meet every nonscalar orbit, and Corollary 2.5 gives full image. In size two use Lemma 6.2.

For $n \geq 4$, write n as a sum of twos and threes, with at least two summands. Use the same free position in all blocks. Intersect the base open conditions and all cross-rectangle determinant opens from Lemma 6.1. This is possible since p is nonidentical on $M_2(K)$. Fix the frozen block diagonal inputs in this intersection.

For a nonscalar target T, Lemma 6.3 permits any diagonal with sum $\operatorname{tr} T$. Choose this diagonal with distinct entries within each block and nonzero sum in each two-block. The excluded conditions are finitely many proper affine hyperplanes, since there are at least two blocks. Each diagonal compression is nonscalar, and every two-compression has nonzero trace. Independently conjugate the diagonal target blocks into the corresponding fixed slice images. These block similarities leave the cross rectangles arbitrary. Solve the diagonal slice equations and every invertible cross-rectangle equation, and assemble their solutions in the one free input. Conjugate all inputs back. Lemma 2.2 supplies the scalar targets. □

7 Reduction of the quartic coefficients

We now prove the trace-nonzero part of Theorem 1.1. In the coefficient arguments, a trace identity is formal, that is, an identity modulo cyclic rotation of words.

If some single-identity reduction has nonzero trace, Vitas's degree-three theorem [12, Theorem B] gives full image for that cubic, hence for p. Indeed, its image is a conjugation-invariant vector space of nonzero trace; it cannot consist of scalars, since its commutator with a new variable would be an identity of degree four, below the minimal identity degree $2n$ for $n \geq 3$. We may assume that all four single-identity reductions are trace zero. These are four independent linear conditions on the 24 coefficients. For a pair $r < s$, make inputs r, s diagonal, write their two vertex tuples as $u = (a, d)$, $v = (b, e)$, and order the remaining inputs increasingly as B, C. Use the local symbols F, G of (2) for this specialization. Set $H(u, v) = -G(v, u, v)$ and define κ_{rs} to be the coefficient of ad in $P(u, v) - H(u, v)$. The coefficient comparison in Appendix A gives

$$\kappa_{12} = \kappa_{34}, \quad \kappa_{13} = \kappa_{24}, \quad \kappa_{14} = \kappa_{23} = -\kappa_{12} - \kappa_{13}. \tag{7}$$

For the same specialization,

$$P(u, v) - H(u, v) = \kappa_{rs}(a - b)(d - e), \quad u = (a, d),\ v = (b, e), \quad H(u, v) = -G(v, u, v), \tag{8}$$

Trace nonvanishing is $(\kappa_{12}, \kappa_{13}) \neq (0, 0)$. Indeed the quotient of the four-constraint coefficient space by the zero-formal-trace space has dimension two, and these functionals give its coordinates, as the cyclic coefficient matrix in Appendix A shows. For a nonzero κ_{rs}, summing the diagonal outputs at $B = E_{12}$, $C = E_{21}$ gives $\kappa_{rs}(a - b)(d - e)$ by (8). Hence trace nonvanishing is detected already in size two.

From (8), the two-cycle determinant is

$$\Delta = \kappa_{rs}(a - b)(d - e)\big(\kappa_{rs}(a - b)(d - e) - P(u, v) - P(v, u)\big).$$

Proposition 3.3 leaves only coefficients satisfying, for every pair,

$$\kappa_{rs} = 0 \quad \text{or} \quad P(u, v) + P(v, u) = \kappa_{rs}(a - b)(d - e). \tag{9}$$

At each label r, at least two incident κ's are nonzero. For such a pair r, s, set the rth diagonal input to the identity in the second equation of (9). The resulting linear polynomial in the two entries of the sth diagonal input is antisymmetric in those entries, hence vanishes when that input is also the identity. Formula (8) gives the same conclusion for the other ordering of B, C. A bilinear polynomial in B, C is a linear combination of BC, CB; its evaluations on $B = E_{12}, C = E_{21}$ detect both coefficients. Thus the double-identity reduction for r, s is zero. A trace-zero cubic has three identity reductions $\mu_1[B, C]$, $\mu_2[A, C]$, $\mu_3[A, B]$, with $\mu_1 - \mu_2 + \mu_3 = 0$, because, for the six coefficients above, trace vanishing is $a + d + e = b + c + f = 0$ and

$$\mu_1 = a + c + d, \quad \mu_2 = a + b + c, \quad \mu_3 = a + b + e.$$

Their indicated alternating sum is $a + d + e = 0$. For each fixed r, at least two incident κ's are nonzero, so at least two identity reductions of the trace-zero cubic $p(X_r = I)$ vanish. The relation among μ_1, μ_2, μ_3 makes the third vanish as well. Hence all double-identity reductions of p are zero, and every single-identity reduction is a proper Lie cubic as in (4).

In (4), the excluded coefficients satisfy $\alpha\beta(\alpha + \beta) = 0$. The cases $\alpha = 0$ and $\beta = 0$ give the two displayed nested commutators; the case $\alpha + \beta = 0$ gives $[[A, B], Y]$ by the Jacobi identity. Thus Lemma 4.2 leaves only the three directions

$$[X_j, [X_k, X_l]], \qquad \{j, k, l\} = \{1, 2, 3, 4\} \setminus \{r\}.$$

For a remaining nonzero reduction at input r, draw an arrow $r \to j$; draw no arrow when that reduction is zero. Fix a free position f and shift the other inputs by scalar multiples

of the identity. Only arrows whose source differs from f contribute. An arrow ending at f gives a nonzero skew coefficient matrix Q in Lemma 4.1; a contributing arrow whose endpoint differs from f gives one nonzero entry, in the row of its outer variable and the column of the remaining frozen variable. Each nonzero reduction coefficient multiplies the independent shift scalar at its source. A nonzero skew contribution already has rank two; adding an independent nonzero single-entry term makes the symmetric part nonzero while preserving a nonzero rank-two minor on a nonempty open set. For two single-entry terms in different rows and columns, the corresponding 2×2 minor is their coefficient product. Their symmetric part is nonzero away from a proper linear condition. Thus each pattern satisfies Lemma 4.1 for some shifts. Lemma 4.2 resolves both patterns.

A path $r \to s \to t$ on three distinct nodes gives, with t free, a skew contribution on the other two frozen positions and an independent single-entry contribution; it is resolved by the preceding test. If $r \leftrightarrow s$ is a mutual pair, any additional arrow starts at one of the other two nodes. If its endpoint is r or s, take the opposite endpoint of the mutual pair as free; if its endpoint is the fourth node, take either endpoint of the mutual pair as free. Each choice gives a skew and a single-entry contribution. A remaining mutual pair is therefore isolated. Without a mutual pair or a three-node path, every arrow ends at a node with zero reduction. Two distinct such endpoints would require disjoint arrows $r \to s$, $t \to u$; taking s free gives the resolved skew and single-entry pattern. The remaining nonempty diagrams are consequently stars with a single zero endpoint, or one isolated mutual pair. Lemma 4.3 resolves every remaining active edge with $\kappa_{rs} \neq 0$. A star with two leaves would force two different incident κ coordinates at its endpoint to vanish. By (7) the three incident coordinates are $\kappa_{12}, \kappa_{13}, -\kappa_{12} - \kappa_{13}$ in some order. Two zero coordinates force all three to vanish, contrary to trace nonvanishing. The empty diagram is proper and is covered by Proposition 5.1.

After permuting variables to A, B, C, D, the remaining case consequently satisfies

$$\begin{aligned} p(C = I) = p(D = I) = 0, \qquad \kappa_{AB} = 0, \\ p(A = I) = \mu[B, [C, D]], \qquad p(B = I) = \nu[A, [C, D]], \end{aligned} \tag{10}$$

with $(\mu, \nu) \neq (0, 0)$. All four cross-pair trace coordinates are nonzero, and their second equations in (9) hold.

For linear operators S, T on the same vector space, put $\{S, T\} = S \circ T + T \circ S$, with composition read from right to left. In the following formulas s_4 means $s_4(A, B, C, D)$ as defined in Section 5. Define the polynomial elements

$$\begin{aligned} P_0 &= A[C, D]B - B[C, D]A, \qquad P_1 = AB[C, D] - [C, D]BA, \\ P_2 &= [A, B][C, D] + [C, D][A, B] + P_0, \\ v_1 &= \tfrac{1}{4}\big(\{\mathrm{ad}_B, \mathrm{ad}_D\}[A, C] - \{\mathrm{ad}_A, \mathrm{ad}_C\}[B, D] \\ &\qquad + \{\mathrm{ad}_B, \mathrm{ad}_C\}[A, D] - \{\mathrm{ad}_A, \mathrm{ad}_D\}[B, C]\big) + \tfrac{1}{2}s_4, \\ v_2 &= \tfrac{1}{4}\big(\{\mathrm{ad}_B, \mathrm{ad}_D\}[A, C] - \{\mathrm{ad}_A, \mathrm{ad}_C\}[B, D] \\ &\qquad - \{\mathrm{ad}_B, \mathrm{ad}_C\}[A, D] + \{\mathrm{ad}_A, \mathrm{ad}_D\}[B, C]\big) - \tfrac{1}{2}s_4. \end{aligned}$$

The conditions just obtained define the six-dimensional space

$$p = z_0 P_0 + z_1 P_1 + z_2 P_2 + a v_1 + b v_2 + d s_4. \tag{11}$$

The equations (10), together with the four cross-pair equations, have rank 18 in the 24-word space, and the six displayed vectors are independent solutions. Appendix A prints their coefficients and a nonzero constraint minor. In these coordinates,

$$\operatorname{tr} p = (-z_0 + z_1 + z_2) \operatorname{tr}([A, B][C, D]). \tag{12}$$

Lemma 7.1. *Suppose identifying two inputs with A, $A^2 = I_n$, and another with $V/2$, $AV = -VA$, leaves the free slice*

$$L(B) = \alpha(ABA)V + \beta V(ABA) + \gamma VB + \delta BV.$$

If

$$\gamma + \delta - \alpha - \beta \neq 0, \qquad (\alpha + \beta + \gamma + \delta)(\gamma^2 - \beta^2)(\delta^2 - \alpha^2) \neq 0,$$

then the original polynomial is surjective in every order $n \geq 3$.

Proof. Choose $V = \operatorname{diag}(v_1, \ldots, v_n)$ with distinct opposite nonzero pairs and, in odd order, one additional zero. Write i^* for the coordinate paired with i, fixing the zero coordinate when present. Let A interchange opposite coordinates and fix the zero coordinate. For a pair of source entries indexed by (i, j) and (i^*, j^*), the coefficient matrix is

$$\begin{pmatrix} s & t \\ -t & -s \end{pmatrix}, \qquad s = \gamma v_i + \delta v_j, \quad t = \beta v_i + \alpha v_j.$$

For a diagonal pair its determinant is $((\alpha+\beta)^2 - (\gamma+\delta)^2)v_i^2$, which is nonzero. Between different nonopposite pairs the determinant is a nonzero polynomial in independent eigenvalue parameters. For a nonzero row and zero column it is $(\beta^2 - \gamma^2)v_i^2$; in the reverse direction it is $(\alpha^2 - \delta^2)v_j^2$. Choose all these determinants nonzero. Set the source entries joining opposite coordinates equal to zero. The singleton diagonal is automatically zero. The image contains the pattern in Lemma 3.2. The diagonal two-cycle map has nonzero trace values, so Lemma 2.2 finishes the proof. □

In (11), identify $C = A$ and set $D = V/2$. The four coefficients of Lemma 7.1 are

$$(\alpha, \beta, \gamma, \delta) = (z_1 + z_2 - b,\ z_1 + z_2 + b,\ z_0 + b,\ z_0 - b).$$

By (12), $z_0 - z_1 - z_2 \neq 0$. After removing nonzero factors, failure of this involution test requires one of

$$z_0 + z_1 + z_2, \qquad z_0 + z_1 + z_2 + 2b, \qquad z_0 + z_1 + z_2 - 2b$$

to vanish. Interchanging the two outer inputs therefore leaves only

$$z_0 + z_1 + z_2 \in \{0, 2b, -2b\}, \qquad z_1 - z_0 - z_2 \in \{0, 2b, -2b\}.$$

The choice $z_0 + z_2 = z_1 = 0$ is proper. The other eight choices are

$$\begin{gathered} z_0 + z_2 = kb, \qquad z_1 = lb, \\ (k, l) \in \{(1, -1), (-1, 1), (-1, -1), (0, -2), (-2, 0), (1, 1), (2, 0), (0, 2)\}. \end{gathered} \tag{13}$$

For $b = 0$ these are again proper. To state the finite slice calculation, define a family of free polynomials, for each listed pair (k, l) and independent parameters $u, a, d \in K$, by

$$q_{k,l}(u, a, d) = lP_1 + kP_2 + v_2 + u(P_0 - P_2) + av_1 + ds_4. \tag{14}$$

For the original coefficients in (11) with $b \neq 0$, the exact relation is

$$p = b\, q_{k,l}(z_0/b, a/b, d/b).$$

Trace nonvanishing for a member of this family is $k + l - 2u \neq 0$.

8 Eight quartic coefficient families

For the family (14), the identity reductions are

$$q_{k,l}(u,a,d)(B=I) = (k+l)[A,[C,D]],$$
$$q_{k,l}(u,a,d)(A=I) = (k-l)[[C,D],B].$$

If $k+l \neq 0$, choose A as the free input. Otherwise choose B; then $k-l \neq 0$ for the listed pairs. In size two take $C = E_{12}$, $D = E_{21}$, so $[C,D] = \mathrm{diag}(1,-1)$. The relevant identity reduction annihilates diagonal matrices and maps E_{12}, E_{21} to nonzero scalar multiples of themselves. Its rank modulo scalars is therefore two. When $k+l-2u \neq 0$, equation (12) also gives nonzero trace in size two: at $A = C = E_{12}$ and $B = D = E_{21}$ it is $2(k+l-2u)$. Thus every trace-nonzero member satisfies the size-two hypotheses of Proposition 6.4.

Lemma 8.1. *For every pair in (13) and every $u, a, d \in K$, some slice of (14) in the selected free position has rank at least seven on $M_3(K)$.*

Proof. Use the matrices

$$H = \mathrm{diag}(2,0,-2), \quad E = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

Freeze the other three positions, in their original order, at

$$(H,E,F), \quad (H+I_3,E,F), \quad (H+2I_3,E,F),$$
$$(H,F,E), \quad (H+I_3,F,E).$$

The commutator relations are $[H,E] = 2E$, $[H,F] = -2F$, $[E,F] = H$, and direct multiplication gives $H^2 + 2(EF+FE) = 8I_3$. For an arbitrary matrix Y, the alternating sum expands to

$$\begin{aligned} s_4(Y,H,E,F) &= [Y,H]H + H[Y,H] + 2[Y,E]F + 2F[Y,E] \\ &\quad + 2[Y,F]E + 2E[Y,F] \\ &= [Y, H^2 + 2(EF+FE)] = 0. \end{aligned}$$

Moving the free input or interchanging E, F changes only the sign. Also $s_4(Y, I_3, E, F) = 0$, since the four placements of the identity in each term have alternating signs. Thus scalar shifts of H preserve the zero slice, proving vanishing at all five tuples. Order the matrix units row by row, assigning index $3(i-1)+j-1$ to E_{ij} for $1 \leq i,j \leq 3$. Except for two indicated minors in Appendix B, use rows and columns $0,\ldots,6$. The resulting five minors for each pair are listed there up to nonzero rational factors. Appendix B proves directly that these five minors have no common zero: on every factor of the first minor, two specified univariate specializations or resultants are coprime. Hence one of the five slices has rank at least seven for every (u,a), independently of d. □

Proof of Theorem 1.1. For a trace-nonzero quartic, the reductions of Section 7 either give surjectivity directly, give a proper quartic covered by Proposition 5.1, or give a nonzero scalar multiple of a member of (14). In the last case Lemma 8.1 and the size-two argument above satisfy Proposition 6.4. Thus every trace-nonzero quartic is surjective for $n \geq 3$. A nonzero trace-zero quartic has image exactly $\mathfrak{sl}_n(K)$ by [13, Theorem 8]. The zero polynomial has zero image. For $n = 2$ apply the established classification [10], allowing $\{0\}, KI_2, \mathfrak{sl}_2(K), M_2(K)$. For $n = 1$ evaluation is scalar multiplication by the sum of the word coefficients, so its image is $\{0\}$ or K. □

A Quartic coefficient matrices

We specify the coefficient matrices used in Section 7. Order the words lexicographically as permutations of A, B, C, D. A polynomial is its column of 24 coefficients. Replacing an input by I deletes that letter in each word and adds coefficients of the resulting words. The single-identity trace equations add coefficients of the two cyclic classes of the remaining three letters. The proper cubic space is the common kernel of its three deletion maps, giving (4).

Let $c_1, \dots, c_6$ be the formal trace coefficients in the classes represented by

$$ABCD,\ ABDC,\ ACBD,\ ACDB,\ ADBC,\ ADCB,$$

in that order. The eight trace equations of the four single-identity reductions have coefficient matrix

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

Its reduced equations are

$$c_1 = c_6, \quad c_2 = -c_5 - c_6, \quad c_3 = c_5, \quad c_4 = -c_5 - c_6.$$

Thus the trace quotient has dimension two. The definitions below give $\kappa_{12} = -c_5$, $\kappa_{13} = -c_6$; in particular these coordinates vanish simultaneously precisely when all c_i vanish. The sign identities in (7) follow from the same coefficient rule.

For a diagonal pair r, s and complementary ordered inputs B, C, let F_{ij}, H_{ij}, $i, j \in \{0, 1\}$, be coefficient functionals on the word space. In a word with B before C, assign that word to F_{ij} when i records whether r lies strictly between B, C and j records the same for s. In a word with C before B, assign its negative to H_{ij} when i, j instead record whether the respective diagonal letter lies outside that interval. Then

$$P = F_{00}ad + F_{01}ae + F_{10}bd + F_{11}be, \quad H = H_{00}ad + H_{01}ae + H_{10}bd + H_{11}be.$$

The common notation in these formulas means evaluation of each coefficient functional on p. Set $\kappa_{rs} = F_{00} - H_{00}$. Modulo the four single-identity trace equations,

$$(F_{ij} - H_{ij})_{ij} = \kappa_{rs}(1, -1, -1, 1),$$

which proves (8). The coefficient equations for the second alternative in (9) are

$$F_{00} + F_{11} = \kappa_{rs}, \qquad F_{01} + F_{10} = -\kappa_{rs}; \tag{15}$$

the other two are duplicates under exchanging the two vertices. This specifies the needed coefficient rows with integer entries.

To verify (11), impose the deletion conditions (10); allow independent scalars μ, ν by requiring the two cubic coefficient vectors to be proportional to their displayed cubic words. Add $\kappa_{AB} = 0$ and (15) for AC, AD, BC, BD. The minors and kernel columns below give rank 18. Their cyclic trace sums give (12).

The six columns, in the lexicographic word order, are as follows.

word	P_0	P_1	P_2	v_1	v_2	s_4
$ABCD$	0	1	1	1	−1	1
$ABDC$	0	−1	−1	0	1	−1
$ACBD$	0	0	0	−1	1	−1
$ACDB$	1	0	1	1	0	1
$ADBC$	0	0	0	0	−1	1
$ADCB$	−1	0	−1	0	0	−1
$BACD$	0	0	−1	−1	0	−1
$BADC$	0	0	1	0	0	1
$BCAD$	0	0	0	1	0	1
$BCDA$	−1	0	−1	−1	1	−1
$BDAC$	0	0	0	0	0	−1
$BDCA$	1	0	1	0	−1	1
$CABD$	0	0	0	0	−1	1
$CADB$	0	0	0	−1	1	−1
$CBAD$	0	0	0	0	0	−1
$CBDA$	0	0	0	1	0	1
$CDAB$	0	0	1	1	−1	1
$CDBA$	0	−1	−1	−1	0	−1
$DABC$	0	0	0	−1	1	−1
$DACB$	0	0	0	0	−1	1
$DBAC$	0	0	0	1	0	1
$DBCA$	0	0	0	0	0	−1
$DCAB$	0	0	−1	0	1	−1
$DCBA$	0	1	1	0	0	1

Their minor in rows $0, 1, 2, 3, 6, 8$ has determinant 1, so they are independent.

Here are explicit indices for the rank-18 calculation. Number the 24 words from zero. Let D_r be the 6×24 deletion matrix for input r, with the six surviving words in lexicographic order. For $r = A, B$, let q_r be the six-entry coefficient vector of $[B, [C, D]]$, respectively $[A, [C, D]]$, whose first entry is one. Form the six rows

$$(D_r)_{j,*} - (q_r)_j (D_r)_{0,*}, \qquad 0 \le j < 6.$$

Follow these twelve rows by the six rows of D_C, the six rows of D_D, the row κ_{AB}, and the two rows in (15) for each of AC, AD, BC, BD, in that order, with all equations moved to the left side. This defines a 33×24 integer matrix. Its submatrix on rows

$$1, 2, 3, 4, 5, 7, 8, 10, 12, 13, 16, 18, 19, 24, 25, 27, 29, 31$$

and columns

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18$$

has determinant 4. Multiplication by the six printed columns gives zero. Thus its rank is both at least 18 and at most $24 - 6 = 18$, establishing (11).

B Rank-seven minors

For each pair (k, l), let $f_0, \ldots, f_4$ denote the five minors in Lemma 8.1, after division by their nonzero rational factors. Table 1 gives these factorizations in the original parameters u, a of (14). The determinant zero sets are unaffected by the removed factors. For $(k, l) =$

$(-1,-1)$, f_0 uses rows $0,1,2,4,5,6,7$; for $(k,l)=(0,-2)$, f_1 uses those rows. All columns and all other rows are $0,\dots,6$.

Table 1: The five minors for each coefficient family.

(k,l)	i	$f_i(u,a)$
$(1,-1)$	0	$(3a-1)(3a+1)(u-2a-3)$ $\cdot(u^2-4u-3a^2+2a+4)$
$(1,-1)$	1	$(a+1)(3a+1)(u-2a-4)$ $\cdot(u^2-4u-3a^2-2a+5)$
$(1,-1)$	2	$(a+1)(3a+5)(u-2a-5)$ $\cdot(u^2-4u-3a^2-6a+4)$
$(1,-1)$	3	$(3a-1)(3a+1)(u+2a-3)$ $\cdot(u^2-4u-3a^2-2a+4)$
$(1,-1)$	4	$(a-1)(3a-1)(u+2a-4)$ $\cdot(u^2-4u-3a^2+2a+5)$
$(-1,1)$	0	$(3a-1)(3a+1)(u-2a+1)$ $\cdot(u^2+4u-3a^2+2a+4)$
$(-1,1)$	1	$(a-1)(3a-1)(u-2a+2)$ $\cdot(u^2+4u-3a^2+6a+1)$
$(-1,1)$	2	$(a-1)(3a-5)(u-2a+3)$ $\cdot(u^2+4u-3a^2+10a-4)$
$(-1,1)$	3	$(3a-1)(3a+1)(u+2a+1)$ $\cdot(u^2+4u-3a^2-2a+4)$
$(-1,1)$	4	$(a+1)(3a+1)(u+2a+2)$ $\cdot(u^2+4u-3a^2-6a+1)$
$(-1,-1)$	0	$(a-1)(3a-1)(3a+1)$ $\cdot(u^2+6u-3a^2-2a+9)$
$(-1,-1)$	1	$(a-1)(3a-1)(u-2a+5)$ $\cdot(u^2+6u-3a^2+2a+10)$
$(-1,-1)$	2	$(a-1)(3a-5)(u-2a+6)$ $\cdot(u^2+6u-3a^2+6a+9)$
$(-1,-1)$	3	$(3a-1)(3a+1)(u+2a+4)$ $\cdot(u^2+6u-3a^2+2a+9)$
$(-1,-1)$	4	$(a+1)(3a+1)(u+2a+5)$ $\cdot(u^2+6u-3a^2-2a+10)$
$(0,-2)$	0	$(3a-1)(3a+1)(u-2a+3)$ $\cdot(u^2+4u-3a^2-2a+4)$
$(0,-2)$	1	$(3a-1)(a-1)^2$ $\cdot(u^2+4u-3a^2+2a+5)$

(k,l)	i	$f_i(u,a)$
$(0,-2)$	2	$(a-1)(3a-5)(u-2a+5)$ $\cdot(u^2+4u-3a^2+6a+4)$
$(0,-2)$	3	$(3a-1)(3a+1)(u+2a+3)$ $\cdot(u^2+4u-3a^2+2a+4)$
$(0,-2)$	4	$(a+1)(3a+1)(u+2a+4)$ $\cdot(u^2+4u-3a^2-2a+5)$
$(-2,0)$	0	$(3a-1)(3a+1)(u-2a+5)$ $\cdot(u^2+8u-3a^2-2a+16)$
$(-2,0)$	1	$(a-1)(3a-1)(u-2a+6)$ $\cdot(u^2+8u-3a^2+2a+17)$
$(-2,0)$	2	$(a-1)(3a-5)(u-2a+7)$ $\cdot(u^2+8u-3a^2+6a+16)$
$(-2,0)$	3	$(3a-1)(3a+1)(u+2a+5)$ $\cdot(u^2+8u-3a^2+2a+16)$
$(-2,0)$	4	$(a+1)(3a+1)(u+2a+6)$ $\cdot(u^2+8u-3a^2-2a+17)$
$(1,1)$	0	$(3a-1)(3a+1)(u-2a-2)$ $\cdot(u^2-6u-3a^2-2a+9)$
$(1,1)$	1	$(a+1)(3a+1)(u-2a-3)$ $\cdot(u^2-6u-3a^2-6a+6)$
$(1,1)$	2	$(a+1)(3a+5)(u-2a-4)$ $\cdot(u^2-6u-3a^2-10a+1)$
$(1,1)$	3	$(3a-1)(3a+1)(u+2a-2)$ $\cdot(u^2-6u-3a^2+2a+9)$
$(1,1)$	4	$(a-1)(3a-1)(u+2a-3)$ $\cdot(u^2-6u-3a^2+6a+6)$
$(2,0)$	0	$(3a-1)(3a+1)(u-2a-3)$ $\cdot(u^2-8u-3a^2-2a+16)$
$(2,0)$	1	$(a+1)(3a+1)(u-2a-4)$ $\cdot(u^2-8u-3a^2-6a+13)$
$(2,0)$	2	$(a+1)(3a+5)(u-2a-5)$ $\cdot(u^2-8u-3a^2-10a+8)$
$(2,0)$	3	$(3a-1)(3a+1)(u+2a-3)$ $\cdot(u^2-8u-3a^2+2a+16)$
$(2,0)$	4	$(a-1)(3a-1)(u+2a-4)$ $\cdot(u^2-8u-3a^2+6a+13)$
$(0,2)$	0	$(3a-1)(3a+1)(u-2a-1)$ $\cdot(u^2-4u-3a^2-2a+4)$

(k,l)	i	$f_i(u,a)$
$(0,2)$	1	$(a+1)(3a+1)(u-2a-2)$ $\cdot(u^2-4u-3a^2-6a+1)$
$(0,2)$	2	$(a+1)(3a+5)(u-2a-3)$ $\cdot(u^2-4u-3a^2-10a-4)$
$(0,2)$	3	$(3a-1)(3a+1)(u+2a-1)$ $\cdot(u^2-4u-3a^2+2a+4)$
$(0,2)$	4	$(a-1)(3a-1)(u+2a-2)$ $\cdot(u^2-4u-3a^2+6a+1)$

For each fixed (k,l), the five minors in Table 1 have no common zero, as follows. For polynomials in one variable over $\mathbb{Q}$, the square-free part means the product of the distinct monic irreducible factors; multiplying it by a nonzero rational number does not affect its roots. Table 2 lists such square-free parts, with rational constants removed.

Fix a pair (k,l) and an irreducible factor g of its first minor f_0. If g depends only on a, it is linear in our table. At its root a_0, Table 2 gives two coprime square-free parts of $f_i(u,a_0)$ and $f_j(u,a_0)$. Thus $g = f_i = f_j = 0$ is impossible.

If g depends on u, regard it and the f_i as polynomials in u over $\mathbb{Q}[a]$. The resultant $\mathrm{Res}_u(g,f_i)$ is the determinant of the Sylvester matrix: for degrees r,s in u, this matrix has s shifted coefficient rows of g and r shifted coefficient rows of f_i, representing multiplication by $1,u,\dots,u^{s-1}$ and $1,u,\dots,u^{r-1}$ respectively. A common finite root in u forces its determinant to vanish, by evaluating those rows at that root. Each such g is monic in u in our table, so specialization in a preserves its positive degree. The two resultants specified in Table 2 have coprime square-free parts. Therefore a common zero of g, f_i, f_j would force a common root of two coprime univariate polynomials in a, a contradiction.

Every factor of f_0 is covered by one row group. In each group the two displayed polynomials have greatest common divisor one over $\mathbb{Q}$, as is checked by their displayed linear and quadratic factors. Coprimality persists over K: the Euclidean algorithm expresses 1 as a rational-polynomial combination of the two polynomials. This proves the assertion in Lemma 8.1. The accompanying supplement additionally verifies the determinant calculations from matrix units and supplies polynomial combinations of the five original minors equal to 1.

Table 2: Elimination on each factor of the first minor.

(k,l)	index	polynomial
$g = 3a-1$; specialize at $\frac{1}{3}$		
$(1,-1)$	1	$(u-2)(3u-14)$
$(1,-1)$	2	$(3u-17)(3u^2-12u+5)$
$g = 3a+1$; specialize at $-\frac{1}{3}$		
$(1,-1)$	2	$(3u-13)(3u^2-12u+17)$
$(1,-1)$	4	$(u-2)(3u-14)$

(k,l)	index	polynomial
$g=-2a+u-3$; resultants in u		
$(1,-1)$	1	$\big(a+1\big)\big(3a+1\big)\big(a^2+2a+2\big)$
$(1,-1)$	4	$\big(a-1\big)\big(3a-1\big)\big(4a-1\big)\big(a^2+6a+2\big)$
$g=-3a^2+2a+u^2-4u+4$; resultants in u		
$(1,-1)$	1	$\big(a+1\big)\big(3a+1\big)\big(4a-1\big)\big(a^2+10a+4\big)$
$(1,-1)$	4	$\big(a-1\big)\big(3a-1\big)\big(a^2-6a+4\big)$
$g=3a-1$; specialize at $\frac{1}{3}$		
$(-1,1)$	2	$\big(3u+7\big)\big(u^2+4u-1\big)$
$(-1,1)$	4	$\big(3u+8\big)\big(3u^2+12u-4\big)$
$g=3a+1$; specialize at $-\frac{1}{3}$		
$(-1,1)$	1	$\big(3u+8\big)\big(3u^2+12u-4\big)$
$(-1,1)$	2	$\big(3u+11\big)\big(3u^2+12u-23\big)$
$g=-2a+u+1$; resultants in u		
$(-1,1)$	1	$\big(a-1\big)\big(3a-1\big)\big(a^2+10a-2\big)$
$(-1,1)$	4	$\big(a+1\big)\big(3a+1\big)\big(4a+1\big)\big(a^2-2a-2\big)$
$g=-3a^2+2a+u^2+4u+4$; resultants in u		
$(-1,1)$	2	$\big(a-1\big)\big(3a-5\big)$
		$\big(a\big)\big(a+1\big)\big(a+2\big)\big(3a+1\big)$
$(-1,1)$	4	$\cdot\big(8a+3\big)$
$g=a-1$; specialize at 1		
$(-1,-1)$	3	$\big(u+2\big)\big(u+4\big)\big(u+6\big)$
$(-1,-1)$	4	$\big(u+1\big)\big(u+5\big)\big(u+7\big)$
$g=3a-1$; specialize at $\frac{1}{3}$		
$(-1,-1)$	2	$\big(3u+16\big)\big(3u^2+18u+32\big)$
$(-1,-1)$	4	$\big(u+3\big)\big(3u+17\big)$
$g=3a+1$; specialize at $-\frac{1}{3}$		
$(-1,-1)$	1	$\big(u+3\big)\big(3u+17\big)$
$(-1,-1)$	2	$\big(3u+20\big)\big(3u^2+18u+20\big)$
$g=-3a^2-2a+u^2+6u+9$; resultants in u		
$(-1,-1)$	1	$\big(a-1\big)\big(3a-1\big)\big(4a+1\big)\big(a^2-10a+4\big)$
$(-1,-1)$	4	$\big(a+1\big)\big(3a+1\big)\big(a^2+6a+4\big)$
$g=3a-1$; specialize at $\frac{1}{3}$		
$(0,-2)$	2	$\big(3u+13\big)\big(3u^2+12u+17\big)$
$(0,-2)$	4	$\big(u+2\big)\big(3u+14\big)$
$g=3a+1$; specialize at $-\frac{1}{3}$		
$(0,-2)$	1	$\big(u+2\big)$
$(0,-2)$	2	$\big(3u+17\big)\big(3u^2+12u+5\big)$

(k,l)	index	polynomial
$g=-2a+u+3$; resultants in u		
$(0,-2)$	1	$\big(a-1\big)\big(3a-1\big)\big(a^2-2a+2\big)$
$(0,-2)$	4	$\big(a+1\big)\big(3a+1\big)\big(4a+1\big)\big(a^2-6a+2\big)$
$g=-3a^2-2a+u^2+4u+4$; resultants in u		
$(0,-2)$	1	$\big(a-1\big)\big(3a-1\big)\big(4a+1\big)$
$(0,-2)$	4	$\big(a+1\big)\big(3a+1\big)\big(a^2+6a+4\big)$
$g=3a-1$; specialize at $\frac{1}{3}$		
$(-2,0)$	2	$\big(3u+19\big)\big(3u^2+24u+53\big)$
$(-2,0)$	4	$\big(u+4\big)\big(3u+20\big)$
$g=3a+1$; specialize at $-\frac{1}{3}$		
$(-2,0)$	1	$\big(u+4\big)\big(3u+20\big)$
$(-2,0)$	2	$\big(3u+23\big)\big(3u^2+24u+41\big)$
$g=-2a+u+5$; resultants in u		
$(-2,0)$	1	$\big(a-1\big)\big(3a-1\big)\big(a^2-2a+2\big)$
$(-2,0)$	4	$\big(a+1\big)\big(3a+1\big)\big(4a+1\big)\big(a^2-6a+2\big)$
$g=-3a^2-2a+u^2+8u+16$; resultants in u		
$(-2,0)$	1	$\big(a-1\big)\big(3a-1\big)\big(4a+1\big)\big(a^2-10a+4\big)$
$(-2,0)$	4	$\big(a+1\big)\big(3a+1\big)\big(a^2+6a+4\big)$
$g=3a-1$; specialize at $\frac{1}{3}$		
$(1,1)$	1	$\big(3u-11\big)\big(3u^2-18u+11\big)$
$(1,1)$	2	$\big(3u-14\big)\big(3u^2-18u-8\big)$
$g=3a+1$; specialize at $-\frac{1}{3}$		
$(1,1)$	2	$\big(3u-10\big)\big(u^2-6u+4\big)$
$(1,1)$	4	$\big(3u-11\big)\big(3u^2-18u+11\big)$
$g=-2a+u-2$; resultants in u		
$(1,1)$	1	$\big(a+1\big)\big(3a+1\big)\big(a^2-10a-2\big)$
$(1,1)$	4	$\big(a-1\big)\big(3a-1\big)\big(4a-1\big)\big(a^2+2a-2\big)$
$g=-3a^2-2a+u^2-6u+9$; resultants in u		
$(1,1)$	2	$\big(a+1\big)\big(3a+5\big)$ $\big(a-2\big)\big(a-1\big)\big(a\big)\big(3a-1\big)$
$(1,1)$	4	$\cdot\big(8a-3\big)$
$g=3a-1$; specialize at $\frac{1}{3}$		
$(2,0)$	1	$\big(3u-14\big)\big(3u^2-24u+32\big)$
$(2,0)$	2	$\big(3u-17\big)\big(3u^2-24u+13\big)$
$g=3a+1$; specialize at $-\frac{1}{3}$		
$(2,0)$	2	$\big(3u-13\big)\big(u^2-8u+11\big)$
$(2,0)$	4	$\big(3u-14\big)\big(3u^2-24u+32\big)$

(k,l)	index	polynomial
$g = -2a + u - 3$; resultants in u		
$(2,0)$	1	$(a+1)(3a+1)(a^2-10a-2)$
$(2,0)$	4	$(a-1)(3a-1)(4a-1)(a^2+2a-2)$
$g = -3a^2 - 2a + u^2 - 8u + 16$; resultants in u		
$(2,0)$	2	$(a+1)(3a+5)$
		$(a-2)(a-1)(a)(3a-1)$
$(2,0)$	4	$\cdot(8a-3)$
$g = 3a - 1$; specialize at $\frac{1}{3}$		
$(0,2)$	1	$(3u-8)(3u^2-12u-4)$
$(0,2)$	2	$(3u-11)(3u^2-12u-23)$
$g = 3a + 1$; specialize at $-\frac{1}{3}$		
$(0,2)$	2	$(3u-7)(u^2-4u-1)$
$(0,2)$	4	$(3u-8)(3u^2-12u-4)$
$g = -2a + u - 1$; resultants in u		
$(0,2)$	1	$(a+1)(3a+1)(a^2-10a-2)$
$(0,2)$	4	$(a-1)(3a-1)(4a-1)(a^2+2a-2)$
$g = -3a^2 - 2a + u^2 - 4u + 4$; resultants in u		
$(0,2)$	2	$(a+1)(3a+5)$
		$(a-2)(a-1)(a)(3a-1)$
$(0,2)$	4	$\cdot(8a-3)$

References

[1] K. Shoda, *Einige Sätze über Matrizen*, Japanese J. Math. **13** (1936), 361–365, doi:10.4099/jjm1924.13.0_361.

[2] A. A. Albert and B. Muckenhoupt, *On matrices of trace zeros*, Michigan Math. J. **4** (1957), 1–3.

[3] S. A. Amitsur and J. Levitzki, *Minimal identities for algebras*, Proc. Amer. Math. Soc. **1** (1950), 449–463.

[4] E. Formanek, *Central polynomials for matrix rings*, J. Algebra **23** (1972), 129–132.

[5] M. Brešar and I. Klep, *Values of noncommutative polynomials, Lie skew-ideals and tracial Nullstellensätze*, Math. Res. Lett. **16** (2009), 605–626; arXiv:0810.1774.

[6] Z. Mesyan, *Polynomials of small degree evaluated on matrices*, Linear Multilinear Algebra **61** (2013), 1487–1495; arXiv:1212.1925.

[7] D. Buzinski and R. Winstanley, *On multilinear polynomials in four variables evaluated on matrices*, Linear Algebra Appl. **439** (2013), 2712–2719; arXiv:1511.06331.

[8] A. Kanel-Belov, S. Malev, and L. Rowen, *The images of multilinear polynomials evaluated on* 3×3 *matrices*, Proc. Amer. Math. Soc. **144** (2016), 7–19; arXiv:1306.4389v3.

[9] A. Kanel-Belov, S. Malev, L. Rowen, and R. Yavich, *Evaluations of noncommutative polynomials on algebras: methods and problems, and the L'vov–Kaplansky conjecture*, arXiv:1909.07785v4.

[10] A. Kanel-Belov, S. Malev, and L. Rowen, *The images of non-commutative polynomials evaluated on* 2×2 *matrices*, Proc. Amer. Math. Soc. **140** (2012), 465–478.

[11] K. Dykema and I. Klep, *Instances of the Kaplansky–Lvov multilinear conjecture for polynomials of degree three*, Linear Algebra Appl. **508** (2016), 272–288.

[12] D. Vitas, *The L'vov–Kaplansky conjecture for polynomials of degree three*, Linear Algebra Appl. **733** (2026), 205–232, doi:10.1016/j.laa.2025.12.016; arXiv:2310.15600v2.

[13] P. S. Fagundes, T. C. de Mello, and P. H. da Silva dos Santos, *The Mesyan conjecture: a restatement and a correction*, arXiv:2111.13698v1.

[14] M. de Bondt, *Mathieu subspaces of codimension less than* n *of* $\mathrm{Mat}_n(K)$, arXiv:1310.7843v2.

[15] P. A. Fillmore, *On similarity and the diagonal of a matrix*, Amer. Math. Monthly **76** (1969), 167–169.

Zhili College, Tsinghua University, Beijing 100084, China
Email address: tangqy24@mails.tsinghua.edu.cn