hexagon:2611.00195v1

The L'vov–Kaplansky conjecture in degrees three and four over arbitrary fields

Qiyue Tang

November 8, 2026

Abstract

We prove the L'vov–Kaplansky conjecture for multilinear polynomials of degrees three and four on matrix algebras over arbitrary fields.

Contents

1 Introduction . . . 1
2 Traceless values of cubics . . . 4
3 Scalar values of cubics . . . 6
4 Cubic projector evaluations . . . 8
5 Cubics over the two-element field . . . 10
6 Matrix order two . . . 12
7 Block sources for quartics . . . 15
8 Proper quartics . . . 17
9 Two-state evaluations of quartics . . . 22
10 Scalar values and auxiliary states . . . 24
11 The remaining quartics over $\mathbf{F}_2$ . . . 27
12 Linear sources for the standard quartic . . . 30
13 Primary blocks over finite fields . . . 41
14 The quartic classification . . . 45
A Coefficient and matrix certificates . . . 46
B Products with a common commutator factor . . . 50
C Finite matrix sources for the standard quartic . . . 55

1 Introduction

1.1 Commutators and the image problem

Let K be a field and let $n \geq 1$ be an integer. The algebra of n-by-n matrices over K is $M_n(K)$, with identity I_n. For $1 \leq i, j \leq n$, the matrix unit E_{ij} has a 1 in position (i, j) and zeros elsewhere. Its size is given by the ambient matrix space. Write $\operatorname{diag}(a_1, \ldots, a_n)$ for the matrix with entries $a_1, \ldots, a_n$ along the diagonal and zeros elsewhere. For a prime power q write $\mathbf{F}_q$ for the field with q elements. We use the ordinary trace $\operatorname{tr} T = \sum_{i=1}^{n} T_{ii}$ and write $\mathfrak{sl}_n(K) = \{T \in M_n(K) : \operatorname{tr} T = 0\}$. For elements of any associative algebra, put $[A, B] = AB - BA$.

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

Shoda treated additive and multiplicative matrix commutators in his 1936 paper [18]. For additive commutators, the equality $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ gives $\operatorname{tr}[A, B] = 0$. Shoda proved that every traceless matrix over a field of characteristic zero is a single commutator [18, Satz 3]. Albert and Muckenhoupt extended this theorem to arbitrary fields [1]. Consider the quadratic polynomial $ax_1x_2 + bx_2x_1$ for $a, b \in K$. If $a + b \neq 0$, setting $x_2 = I_n$ gives the whole algebra. If $a + b = 0$ and $a \neq 0$, the image is $\mathfrak{sl}_n(K)$. The zero polynomial gives $\{0\}$.

The free associative algebra $K\langle x_1, \ldots, x_d\rangle$ has the words in the generators as a basis, with concatenation as multiplication and the empty word as unit. A multilinear polynomial of degree d in this paper has the form

$$p(x_1, \ldots, x_d) = \sum_{\sigma \in S_d} a_\sigma x_{\sigma(1)} \cdots x_{\sigma(d)}, \qquad a_\sigma \in K,$$

where S_d is the group of permutations of $\{1, \ldots, d\}$. The zero polynomial is included by allowing all coefficients to be zero. Write $p(M_n(K))$ for the set of all $p(A_1, \ldots, A_d)$ with $A_i \in M_n(K)$. The standard polynomial s_d is the sum in which a_σ is the sign of σ.

Two multilinear words are cyclically equivalent if one is obtained from the other by moving an initial segment to the end. The formal trace of p is the list of coefficient sums over these classes. The evaluated trace is the function sending $(A_1, \ldots, A_d)$ to $\operatorname{tr} p(A_1, \ldots, A_d)$. We call a polynomial traceful in a given matrix order if the evaluated trace is nonzero as a function there. A proper polynomial is one whose substitution of the unit for any one generator is zero in the free algebra.

Polynomial identity theory gives the zero and scalar images. A polynomial identity has zero image. Amitsur and Levitzki proved that s_{2n} is an identity of $M_n(K)$ and that $2n$ is the least degree of a nonzero identity [2]. A central polynomial takes scalar values and has at least one nonzero value. Kaplansky posed the existence problem for such polynomials in his report from the 1956 Shelter Island conference [12]. Formanek constructed central polynomials for matrix rings [9]. For a multilinear central polynomial, scaling one input gives the entire scalar line KI_n.

I. V. L'vov asked whether $p(M_n(K))$ is a vector subspace for every field K, multilinear polynomial p and matrix order n [8, Problem 1.98]. Kanel-Belov, Malev and Rowen discuss the attribution to Kaplansky and the formulation by L'vov [11, Introduction]. A multilinear image is closed under scalar multiplication and simultaneous conjugation. The question is whether a sum of two values can again be attained by one tuple. The linear span of the image is a Lie ideal, meaning a subspace closed under taking commutators with arbitrary matrices. Indeed, the product rule gives

$$[p(A_1, \ldots, A_d), B] = \sum_{i=1}^{d} p(A_1, \ldots, A_{i-1}, [A_i, B], A_{i+1}, \ldots, A_d).$$

Brešar and Klep proved the corresponding Lie-ideal result for general polynomial values over infinite fields [3, Theorem 2.3].

1.2 Results at fixed matrix order and degree

For matrices of order two, Kanel-Belov, Malev and Rowen proved the four-space classification over quadratically closed fields for all multilinear degrees [11, Theorem 2]. Malev obtained a partial result over arbitrary fields and the complete classification over real closed fields [13]. Gonzales Gargate and de Mello gave a grading proof of the traceless-containment alternative in characteristic different from two [10, Theorem 5.2 of the preprint].

Fixing the degree, Mesyan proved that the image of a nonzero multilinear polynomial of degree at most three contains $\mathfrak{sl}_n(K)$ when K has at least n elements [14, Theorem 13]. In

degree three and order $n \geq 2$, the sums of coefficients in the two cyclic classes determine whether all values have zero trace [14, Lemma 14]. Dykema and Klep proved the full cubic classification over $\mathbb{C}$ in every even matrix order and in odd orders less than 17 [5, Theorem 1.2]. Vitas proved it in all matrix orders over algebraically closed fields of characteristic zero [22, Theorem B]. His Theorem 3.3 also treats certain positive characteristics over an algebraically closed extension, under a condition on roots of unity in the coefficient field.

In degree four, Buzinski and Winstanley treated traceless containment over algebraically closed fields of characteristic zero by reducing to proper polynomials [4]. Fagundes, de Mello and dos Santos corrected a lemma in that argument and established the containment [6, Theorem 8]. For a nonzero quartic, the containment determines the image when every value is traceless. When the trace function is nonzero, the remaining targets have nonzero trace.

In our earlier Hexagon preprint [20, Theorem 1.1], we obtained the full quartic classification over algebraically closed fields of characteristic zero. In orders $n \geq 3$, the image is zero for the zero polynomial, $\mathfrak{sl}_n(K)$ for a nonzero polynomial with identically zero trace, and $M_n(K)$ when some value has nonzero trace. Our aiXiv preprint [21, Theorem 1] treats degrees five and six over the same fields. In orders $n \geq 3$ it gives the corresponding classification by identities and trace. The only nonzero identities in these orders are the nonzero multiples of s_6 in order three.

Two other Hexagon preprints concern these images over algebraically closed fields of characteristic zero. Pan [17, Theorem A and Corollary 8.12] obtains the classification in degrees four through six, deriving degrees four and five from the sextic case by product substitution. His quartic classification agrees with that in [20, Theorem 1.1]. Miah [15, Theorem 1.1] determines the images of the standard polynomials in every degree and matrix order. His even-degree argument develops the method used for the standard sextic in [21, Appendix A].

Passing to arbitrary fields requires different constructions. Over a general field, a rational canonical block can have an irreducible characteristic polynomial of degree greater than one. If the characteristic divides n, every scalar matrix has trace zero. In a small finite field, a nonzero polynomial in a parameter can vanish at all available values.

1.3 The classification

Lemma 1.1. *For $1 \leq d < 2n$, a nonzero multilinear polynomial of degree d has a nonzero value on $M_n(K)$. For $1 \leq d \leq 2n$, the evaluated trace vanishes identically if and only if the formal trace is zero.*

Proof. For the first assertion choose a word with nonzero coefficient. Assign its letters, in their order in that word, to the first d units in the list

$$E_{11}, E_{12}, E_{22}, E_{23}, \ldots, E_{n-1,n}, E_{nn}.$$

The list consists of E_{ii} for $1 \leq i \leq n$, with $E_{i,i+1}$ immediately after E_{ii} for $i < n$. A product of these units is nonzero exactly when the successive column and row indices agree. The forward edges $i \to i+1$ must therefore occur in increasing order, and each loop must occur during the unique visit to its vertex. This forces the chosen word order. Its product is a matrix unit, and all other words give zero.

For the trace assertion, rotate each word so that x_d is last. Its trace after evaluation is unchanged. The sum of the rotated words has the form $q(x_1, \ldots, x_{d-1})x_d$, where the coefficients of q are the cyclic sums of p. The form $(U, V) \mapsto \operatorname{tr}(UV)$ on $M_n(K)$ is nondegenerate, since $\operatorname{tr}(UE_{ji}) = U_{ij}$. Hence identically zero evaluated trace forces q to vanish on $M_n(K)$.

For $d > 1$ the first assertion applies to q because $d - 1 < 2n$. For $d = 1$ it is the scalar coefficient of x_1 and vanishes by the same pairing. Thus the formal trace is zero. The converse follows directly from the rotations. □

Theorem 1.2. *Let K be any field, let $n \geq 1$, and let p be a multilinear polynomial of degree three or four over K. Its image on $M_n(K)$ is one of*

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

For $n \geq 3$ it is zero if $p = 0$, equals $\mathfrak{sl}_n(K)$ if $p \neq 0$ has zero formal trace, and equals $M_n(K)$ otherwise.

In degree three, the three-way classification holds already for $n = 2$. For a quartic the scalar alternative occurs, for example, in characteristic different from two in order two. For traceless $U, V \in M_2(K)$, the Cayley–Hamilton identity gives $U^2 = -(\det U)I_2$, and likewise for V and $U + V$. Subtracting the identities for U and V from the one for $U + V$ gives

$$UV + VU = (\det U + \det V - \det(U + V))I_2.$$

Apply this to the two commutators in

$$[x_1, x_2][x_3, x_4] + [x_3, x_4][x_1, x_2].$$

Taking both commutators equal to $[E_{12}, E_{21}] = \mathrm{diag}(1, -1)$ gives $2I_2$.

For the cubic case write

$$f(x, y, z) = axyz + bxzy + cyxz + dyzx + ezxy + gzyx, \qquad a, b, c, d, e, g \in K. \tag{1}$$

Theorem 1.3. *Let K be any field, let f be as in (1), and set $S = a + b + c + d + e + g$ and $u = a + d + e$. For $n \geq 2$,*

$$f(M_n(K)) = \begin{cases} \{0\}, & f = 0, \\ \mathfrak{sl}_n(K), & f \neq 0,\ S = u = 0, \\ M_n(K), & \text{otherwise.} \end{cases}$$

For $n = 1$, the image is $\{0\}$ if $S = 0$ and K otherwise.

For traceless cubic targets, we prescribe trace pairings against powers of a companion matrix, adapting Stasinski's construction [19, Section 4]. For traceful cubics, substituting a scalar shift of the idempotent $\mathrm{diag}(1, \ldots, 1, 0)$ separates the equations on the two diagonal and the two off-diagonal blocks. The fields of orders two, three and four require separate evaluations. In degree four, substitutions by the identity and repeated inputs reduce the coefficient cases. The proof for s_4 constructs values on companion blocks and joins them by rectangular linear maps.

2 Traceless values of cubics

For $m \geq 1$ and $A \in M_m(K)$, write ad_A for the linear map $B \mapsto [A, B]$ on $M_m(K)$. A matrix A is cyclic if there is a column v for which $v, Av, \ldots, A^{m-1}v$ form a basis. In this case its centralizer, the space of matrices commuting with it, is the span of $I_m, A, \ldots, A^{m-1}$. Indeed, a matrix commuting with A is determined by its value on v, and a polynomial in A of degree less than m can assign that value arbitrarily. If U commutes with A, then $\mathrm{tr}(U[A, B]) = \mathrm{tr}((UA - AU)B) = 0$. By rank-nullity, the image of ad_A has codimension equal to the dimension of the centralizer. For cyclic A, the image is therefore the orthogonal complement of its centralizer, giving

$$T \in \mathrm{ad}_A(M_m(K)) \quad \Longleftrightarrow \quad \mathrm{tr}(TA^j) = 0 \quad (0 \leq j < m). \tag{2}$$

Lemma 2.1. *Over any field, every traceless matrix is a commutator $[A, B]$. If the target is nonscalar, it is also a double commutator $[A, [A, B]]$. For a nonscalar target of size at least three, both representations can be obtained with a traceless A whose characteristic polynomial is separable.*

Proof. If a matrix T is nonscalar, there is a linear functional ℓ such that ℓ and ℓT are independent. Otherwise every nonzero functional would be an eigenvector for composition with T. Applying this to a basis and the sums of pairs of basis functionals would force all its eigenvalues to coincide. The same argument on column vectors gives a column v with v, Tv independent whenever T is nonscalar.

Let T be nonscalar and traceless of size $n \geq 3$, and put $m = n - 1$. For $1 \leq i \leq m$ write e_i for the ith standard column of K^m, and use a superscript T for transpose. Extend $\ell, \ell T$ to a basis of the dual space and take the corresponding dual basis of K^n. Since the first coordinate of T in this basis is ℓT, the first row has a 1 in position 2 and zeros elsewhere. Thus

$$T = \begin{pmatrix} 0 & e_1^{\mathsf{T}} \\ v & D \end{pmatrix},$$

with a column v and a square block D. Choose a scalar a_0 and a monic polynomial $F(t)$ of degree m. If m is nonzero in K, take $a_0 = 0$ and $F(t) = t^m - 1$. If $m = 0$ in K and $m \geq 3$, take $a_0 = 0$ and $F(t) = t^m + t + 1$. For $m = 2$ in characteristic two, take $a_0 = 1$ and $F(t) = t^2 + t + 1$. The derivatives are mt^{m-1}, 1, and 1, respectively. In the first case $F(0) = -1$, so F and its derivative have no common root. In each case $F(a_0) \neq 0$, and the coefficient of t^{m-1} is a_0. Let P be its row companion matrix, whose first $m - 1$ rows are $e_2^{\mathsf{T}}, \ldots, e_m^{\mathsf{T}}$ and whose last row lists the negatives of the coefficients in increasing degree below m. In particular, $e_1^{\mathsf{T}} P^j = e_{j+1}^{\mathsf{T}}$ for $0 \leq j < m$.

For a column $x \in K^m$ let

$$A = \begin{pmatrix} a_0 & 0 \\ x & P \end{pmatrix}.$$

Conjugating $\operatorname{diag}(a_0, P)$ by $\left(\begin{smallmatrix} 1 & 0 \\ -(P - a_0 I_m)^{-1} x & I_m \end{smallmatrix}\right)$ gives A. The trace of A is $a_0 - a_0 = 0$, and its characteristic polynomial $(t - a_0)F(t)$ is separable. Multiplication by A on the left shows inductively that the lower-left block of A^j is $\sum_{i=0}^{j-1} a_0^{j-1-i} P^i x$. For $1 \leq j \leq m$,

$$\operatorname{tr}(TA^j) = e_1^{\mathsf{T}} \left(\sum_{i=0}^{j-1} a_0^{j-1-i} P^i \right) x + \operatorname{tr}(DP^j).$$

The coefficient of x_j is one and only $x_1, \ldots, x_j$ occur. Hence the m equations $\operatorname{tr}(TA^j) = 0$ determine x successively. The matrix A has n distinct eigenvalues over a splitting field. Its minimal and characteristic polynomials coincide, so its rational canonical form has one companion block. This proves the commutator assertion by (2). Over a splitting field ad_A is diagonal in matrix units relative to an eigenbasis of A. Thus ad_A and ad_A^2 have equal rank there, and also over K because scalar extension preserves rank. Their images over K are nested, so they are equal. This gives the double commutator.

For nonscalar traceless T of size two, choose a basis v, Tv. The first basis vector maps to the second, and $T^2 = -(\det T) I_2$ since $\operatorname{tr} T = 0$. Hence the diagonal of T in this basis is zero. Taking $A = E_{11}$, the map ad_A is invertible on the off-diagonal subspace and ad_A^2 is the identity there. For a nonzero traceless scalar, the characteristic is a prime p dividing the matrix order. In size p, take

$$A = \sum_{j=1}^{p-1} E_{j,j+1}, \qquad B = \sum_{j=1}^{p-1} j E_{j+1,j}.$$

The products are $AB = \mathrm{diag}(1, 2, \ldots, p-1, 0)$ and $BA = \mathrm{diag}(0, 1, \ldots, p-1)$. Their difference is I_p in characteristic p. Scaling one input and using block-diagonal copies gives every traceless scalar. The zero matrix is itself a commutator. □

Proposition 2.2. *If f is a nonzero multilinear cubic, then $\mathfrak{sl}_n(K) \subseteq f(M_n(K))$ for every $n \geq 2$.*

Proof. If replacing one variable by the unit gives a nonzero polynomial, write it as $rUV+sVU$ in the two remaining inputs U, V, with $r, s \in K$. If $r + s \neq 0$, its image is full by setting $V = I_n$. If $r + s = 0$, it is a nonzero multiple of a commutator and Lemma 2.1 applies.

If all three replacements give zero polynomials, the coefficient equations are

variable replaced	coefficients in the two remaining word orders	
x	$a + c + d = 0$	$b + e + g = 0$
y	$a + b + c = 0$	$d + e + g = 0$
z	$a + b + e = 0$	$c + d + g = 0.$

Comparing the first and third equations of the first column with its second equation gives $d = b$ and $e = c$. The first column also gives $a + b + c = 0$, and the first row of the second column gives $g = -b - c = a$. Expanding the two brackets gives

$$f(x, y, z) = a[[x, y], z] + b[[x, z], y].$$

If $a \neq 0$, setting $x = z = A$ gives $-a[A, [A, y]]$. If $a = 0$, then $b \neq 0$, and setting $x = y = A$ gives $-b[A, [A, z]]$. These substitutions and Lemma 2.1 give all nonscalar traceless matrices.

To obtain a nonzero traceless scalar in odd characteristic p, use the size-p pair A, B from Lemma 2.1. The identity $[A, B^2] = [A, B]B + B[A, B] = 2B$ gives $\mathrm{ad}_A^2(B^2/2) = I_p$. The two substitutions above, scaling, and block-diagonal repetition give the required scalar. In characteristic two the inputs (E_{12}, E_{11}, E_{21}) give $[[E_{12}, E_{11}], E_{21}] = -[E_{12}, E_{21}] = I_2$ and $[[E_{12}, E_{21}], E_{11}] = [I_2, E_{11}] = 0$. Interchanging y and z interchanges the two double commutators. One of these evaluations therefore gives a nonzero multiple of I_2, which yields every traceless scalar by scaling and block-diagonal repetition. □

3 Scalar values of cubics

Cyclicity of trace gives

$$\mathrm{tr}\, f(X, Y, Z) = u\ \mathrm{tr}(XYZ) + (S - u)\ \mathrm{tr}(XZY). \tag{3}$$

The traces on the right can be tested separately using (E_{11}, E_{12}, E_{21}) and interchanging the last two entries. Thus the trace vanishes on every value in size at least two precisely when $S = u = 0$. When $S \neq 0$, one already obtains every target by setting two inputs to the identity. We treat $S = 0$ and $u \neq 0$ from now on. We may divide all coefficients by u whenever useful, since scaling the polynomial by a nonzero scalar leaves its image unchanged. Permuting the input labels also preserves the image and changes u at most by sign.

Proposition 3.1. *If $|K| > 2$, $S = 0$, and $u \neq 0$, then $f(M_n(K))$ contains every scalar matrix for $n \geq 2$.*

Proof. Normalize $u = 1$. First take $n = 2$. In characteristic different from two put

$$X = \mathrm{diag}(1, -1), \qquad Y = E_{12} + E_{21}, \qquad Z = XY.$$

Here $X^2 = Y^2 = I_2$ and $XY = -YX$, so $Z = XY$ anticommutes with X, Y and $XYZ = Z^2 = -I_2$. The products in the cyclic class of xyz equal $-I_2$, and those in the other class equal I_2. The coefficient sums 1 and -1 give $-2I_2$, which scales to I_2. Replacing x, y, or z by the unit gives a commutator with coefficient, in the remaining variables in their original order, respectively

$$\mu_x = a + c + d, \qquad \mu_y = a + b + c, \qquad \mu_z = a + b + e. \tag{4}$$

The identity $\mu_x - \mu_y + \mu_z = u$ ensures that at least one is nonzero. In characteristic two, $[E_{12}, E_{21}] = I_2$, so that specialization gives I_2 after scaling.

For size three first permute the variables to arrange $(c, e) \neq (0, 0)$, then normalize $u = 1$ again if necessary. The three pairs of reverse words are indexed by their middle variable, and some pair has a nonzero coefficient. Let $J = E_{12} + E_{23} + E_{31}$ and put

$$X = \operatorname{diag}(0, 1, t), \qquad Y = J, \qquad Z = \operatorname{diag}(z_0, z_1, z_2)J^{-1}$$

for $t, z_0, z_1, z_2 \in K$. Each word gives a diagonal matrix, with diagonal entries

word	diagonal entries		
$xyz,\ yzx$	0	z_2	tz_0
$xzy,\ zyx$	0	z_1	tz_2
yxz	z_1	tz_2	0
zxy	tz_0	0	z_2.

Using $a + d = 1 - e$ and $b + g = -1 - c$, the matrix of these diagonal entries for f as a linear map of $(z_0, z_1, z_2)^{\mathsf{T}}$ is

$$\begin{pmatrix} et & c & 0 \\ 0 & -(1+c) & 1-e+ct \\ (1-e)t & 0 & e-(1+c)t \end{pmatrix}.$$

The two nonzero permutation terms in its determinant give

$$-et(1+c)\big(e-(1+c)t\big) + ct(1-e)(1-e+ct) = t\big(c - 2ce - e^2 + (e + 2ce + c^2)t\big).$$

If the expression in parentheses is a nonzero polynomial in t, choose a nonzero t at which it is nonzero. This is possible since $|K| > 2$. Then solve for diagonal entries $(1, 1, 1)$.

Otherwise the two coefficients of that expression vanish. Their sum is $(c + e)(1 + c - e)$. If $c + e = 0$, the constant coefficient becomes $c(c + 1)$, giving $(c, e) = (0, 0)$ or $(-1, 1)$. Our choice of variable order leaves $(c, e) = (-1, 1)$. For $t = -1$ the three rows give $-z_0 - z_1$, z_2 and z_2, respectively, so take $(z_0, z_1, z_2) = (-1, 0, 1)$. If $1 + c - e = 0$, substituting $e = c + 1$ into the vanishing constant coefficient gives $3c^2 + 3c + 1 = 0$, or $c^2 + ce + e^2 = 0$. Neither $c = 0$ nor $c = -1$ satisfies $3c^2 + 3c + 1 = 0$. Take $t = 1$ and $(z_0, z_1, z_2) = (-1/c, -1/e, 0)$. The rows give

$$-e/c - c/e = 1, \qquad (1 + c)/e = 1, \qquad -(1 - e)/c = 1.$$

Every integer $n \geq 2$ is a sum of twos and threes. After undoing any variable permutation and scaling one input to undo the normalization in each block, use the size-two and size-three inputs block-diagonally to attain I_n, and scale one input to attain an arbitrary scalar matrix. □

4 Cubic projector evaluations

Lemma 4.1. *If $|K| > 2$, then every traceless $T \in M_m(K)$, $m \geq 1$, is a commutator $[L, P]$ with L invertible.*

Proof. For $m \geq 3$ put T in rational canonical form, with companion blocks having ones on the first subdiagonal and the negatives of the defining coefficients in the last column. Thus T has no entries below its first subdiagonal, and the nonzero entries on that subdiagonal are ones. If there is exactly one such entry, there is one invariant factor of degree two and all other invariant factors have degree one. At least one linear factor divides the quadratic invariant factor, so the characteristic polynomial splits and T is triangularizable. In this case replace the rational canonical form by an upper triangular form. If there are no subdiagonal ones, T is already upper triangular.

Choose a strictly upper bidiagonal N with every first-superdiagonal entry nonzero. If the form of T has at least two subdiagonal ones, choose the corresponding entries of N to have sum zero. A list of any even length can be formed from pairs $(1, -1)$. For odd length at least three, use a triple $(1, \eta, -1-\eta)$ with $\eta \notin \{0, -1\}$ and fill the remaining positions with pairs. At all other positions take the entry of N to be 1. This gives $\operatorname{tr}(TN) = 0$. For $j \geq 2$, the entries of N^j are supported on the jth superdiagonal, so $\operatorname{tr}(TN^j) = 0$. The last standard column and its first $m-1$ iterates under N are nonzero multiples of the standard columns in reverse order. Thus N is cyclic, and (2) gives $T = [N, P] = [I_m + N, P]$.

For $m = 2$ and T nonscalar, use the zero-diagonal form from Lemma 2.1. A diagonal matrix with two distinct nonzero entries gives an invertible L for which ad_L maps the off-diagonal subspace onto itself. The only nonzero traceless scalars in size two occur in characteristic two, where $[I_2 + E_{12}, \lambda E_{21}] = \lambda I_2$. The zero target in any size is $[I_m, 0]$, which also covers $m = 1$. □

Proposition 4.2. *Suppose $|K| > 2$, $S = 0$, and $u \neq 0$. Put $t = a + b + c$ and $s = d + e$. If there is $\lambda \in K$ with*

$$\begin{aligned} t(1+\lambda) &\neq 0, \qquad & s - \lambda t &\neq 0, \\ (a + c + \lambda t, d - \lambda t) &\neq (0,0), \qquad & (e + g - \lambda t, b + \lambda t) &\neq (0,0), \end{aligned} \tag{5}$$

then $f(M_n(K))$ contains every nonscalar matrix for all $n \geq 2$.

Proof. Set $m = n - 1$. Use matrices of the form

$$X = \begin{pmatrix} A & 0 \\ C & 1 \end{pmatrix}, \qquad Y = \operatorname{diag}(I_m, 0) + \lambda I_n, \qquad Z = \begin{pmatrix} P & Q \\ R & 0 \end{pmatrix}, \tag{6}$$

where $A, P \in M_m(K)$, C, R are rows, and Q is a column of length m. For $\lambda = 0$, the word products have the blocks

word	$(1,1)$	$(1,2)$	$(2,1)$	$(2,2)$
xyz	AP	AQ	CP	CQ
xzy	AP	0	$CP + R$	0
yxz	AP	AQ	0	0
yzx	$PA + QC$	Q	0	0
zxy	$PA + QC$	0	RA	0
zyx	PA	0	RA	0.

(7)

The λ term is $\lambda t[X,Z]$, by (4). Thus the blocks f_{ij} of $f(X,Y,Z)$ are

$$\begin{aligned} f_{11} &= t(1+\lambda)[A,P] + (s-\lambda t)QC, \\ f_{12} &= ((a+c+\lambda t)A + (d-\lambda t)I_m)Q, \\ f_{21} &= (a+b+\lambda t)CP + R((e+g-\lambda t)A + (b+\lambda t)I_m), \\ f_{22} &= (a+\lambda t)CQ. \end{aligned} \tag{8}$$

For a nonscalar target W choose v, Wv independent and a functional ℓ satisfying $\ell(v) = 1$ and

$$\ell(Wv) = \frac{a+\lambda t}{u} \operatorname{tr} W.$$

In the decomposition $K^n = \ker \ell \oplus Kv$, write $W = \begin{pmatrix} T & U \\ V & w \end{pmatrix}$. Then $U \neq 0$ and $\operatorname{tr} T = (s-\lambda t)\operatorname{tr} W/u$. Extend U to an ordered basis of K^m and let N send each basis column to the next and the last to zero. Then N is nilpotent and U is cyclic for N. Put $A = \tau I_m + N$. The last two conditions of (5) allow a choice of τ making both linear matrix factors in (8) invertible. Indeed, each factor is a scalar multiple of I_m plus a multiple of N, and its scalar part excludes at most one τ. Take Q to be the inverse of the upper-right factor applied to U. It is cyclic for A, since this factor is an invertible polynomial in A.

Choose C by the m equations

$$CA^jQ = \frac{\operatorname{tr}(TA^j)}{s-\lambda t} \qquad (0 \le j < m).$$

The equations prescribe the values of C on the basis $(A^jQ)_{0\le j<m}$. Since $\operatorname{tr}(QCA^j) = CA^jQ$, the right side of the following equation is orthogonal to every A^j for $0 \le j < m$. By (2) there is a P with

$$t(1+\lambda)[A,P] = T - (s-\lambda t)QC.$$

Solve for R in the lower-left block of (8) by multiplying on the right by the inverse of its linear matrix factor. The lower-right block is $(a+\lambda t)\operatorname{tr} T/(s-\lambda t) = w$. This constructs W in the chosen basis, and conjugation returns to the original basis. □

The identity $\mu_x - \mu_y + \mu_z = u$ in (4) ensures a nonzero unit specialization. Permute the variables so that $t = \mu_y \neq 0$. The first two conditions in (5) exclude only -1 and s/t. The sums of the entries of the two ordered pairs in the remaining conditions are μ_x and $-\mu_x$. Each pair excludes at most one additional value, namely d/t or $-b/t$, respectively, and neither excludes anything when $\mu_x \neq 0$. This proves coverage when $|K| \ge 5$. If at least two unit specializations are nonzero, place them at x, y to obtain coverage also over fields with three or four elements.

4.1 One nonzero unit specialization

The remaining case over fields with three or four elements has exactly one nonzero unit specialization. Place it at y and normalize $u = 1$. The equations $\mu_x = \mu_z = 0$ give $d = -a-c$ and $e = -a-b$. Substitution in $u = a+d+e = 1$ gives $t = a+b+c = -1$. Then $S = 0$ gives $g = a$. The coefficient tuple is

$$(a,b,c,d,e,g) = (a, b, -1-a-b, 1+b, -a-b, a). \tag{9}$$

The values of λ excluded by (5) are

$$-1, \quad a-1, \quad -1-b, \quad b. \tag{10}$$

Swapping x, z sends the coefficients of xyz, xzy to g, e, respectively, and interchanges the two cyclic sums $1, -1$. Multiplication by -1 restores $u = 1$ and sends (a, b) to $(-a, a + b)$.

Over $\mathbf{F}_3$, if $b = 0$ or 2, the last two entries of (10) are 0 and 2, so exclusion of all three elements requires $a = 2$. If $b = 1$, the last two entries both equal 1, so exclusion of all three requires $a = 1$. The pair $(2, 0)$ becomes $(1, 2)$ under the swap and then admits a parameter. The pairs $(1, 1)$ and $(2, 2)$ are interchanged by the swap. For $(2, 2)$ the coefficients in (9) are $(-1, -1, 1, 0, 2, -1)$ in $\mathbf{F}_3$. With $\lambda = 1$, (8) becomes

$$f(X, Y, Z) = \begin{pmatrix} [A, P] & (I_m - A)Q \\ R(I_m - A) & CQ \end{pmatrix}.$$

Choose the target decomposition as in Proposition 4.2, but impose $\ell(Wv) = \operatorname{tr} W$. Then $\operatorname{tr} T = 0$, $U \neq 0$, and $w = \operatorname{tr} W$. By Lemma 4.1 write $T = [L, P_0]$ with L invertible. Set $A = I_m - L$, $P = -P_0$, $Q = L^{-1}U$, and $R = VL^{-1}$. Take any row C with $CQ = w$, which exists since $Q \neq 0$. These inputs give W.

Over $\mathbf{F}_4$, (10) covers the field only when $a = 1$ and $b \notin \mathbf{F}_2$. In fact $b, b + 1$ give just $\mathbf{F}_2$ if $b \in \mathbf{F}_2$, and otherwise they give its complement. In the latter case $1, a + 1$ must supply the two elements of $\mathbf{F}_2$. The coefficients are $(1, b, b, b + 1, b + 1, 1)$, with $b, b + 1 \neq 0$. With $\lambda = 0$, (8) gives

$$f(X, Y, Z) = \begin{pmatrix} [A, P] & (b + 1)(A + I_m)Q \\ (b + 1)CP + bR(A + I_m) & CQ \end{pmatrix}.$$

Use the same target decomposition as over $\mathbf{F}_3$ and write $T = [L, P]$ with L invertible. Set $A = L + I_m$, $Q = (b+1)^{-1}L^{-1}U$, choose C with $CQ = w$, and put $R = b^{-1}(V - (b+1)CP)L^{-1}$. These inputs give W.

5 Cubics over the two-element field

Proposition 5.1. *A multilinear cubic over $\mathbf{F}_2$ with coefficient sum zero and nonzero cyclic coefficient sum is surjective on $M_n(\mathbf{F}_2)$ for every $n \geq 2$.*

Proof. Each of the two cyclic word classes contains an odd number of terms of the polynomial. Renaming the inputs and reversing every word preserve surjectivity, the latter by transposing all inputs and the output. The support is the set of words with nonzero coefficient. In a two-word support, rename the variables to make the word from one cyclic class xyz. The other word is yxz, xzy or zyx. The first gives $[x, y]z$, and the second gives the same form after reversing both words and renaming the inputs. The third gives $xyz - zyx$. A four-word support consists of one full cyclic class and a singleton of the opposite class. Renaming the variables makes the full class the class of xzy, and cyclic renaming is transitive on the opposite class. Thus there is a single representative $h = xzy + yxz + zxy + zyx$. The six-word support gives s_3, whose signs are all one in characteristic two.

For $[x, y]z$, choose an invertible trace-zero matrix, for instance a cyclic permutation matrix of order n. By the commutator assertion in Lemma 2.1, it is some $[x, y]$. Varying z gives every target. For $xyz - zyx$ and a target W, choose invertible P, Q such that $PWQ = \operatorname{diag}(I_r, 0)$, where r is the rank of W. If J is an invertible cyclic permutation matrix of order n, take $A^{-1} = QJP$. Then

$$\operatorname{tr}(WA^{-1}) = \operatorname{tr}(PWQJ) = 0.$$

Write $WA^{-1} = [D, T]$. The inputs $(A, A^{-1}D, TA)$ give $DTA - TDA = [D, T]A = W$.

For h, take the fixed pairs of square matrices

$$Y_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Z_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad Y_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

When the second and third inputs are block diagonal with these blocks, the map on an i-by-j block X of the first input is

$$L_{ij}(X) = XZ_jY_j + Y_iXZ_j + Z_iXY_j + Z_iY_iX.$$

Use the matrix units in row order in both source and target. Multiplication gives the following rows for the three indicated matrices of L_{ij}, writing each row as a string of entries:

ij	Rows
22	0010, 1000, 0001, 0100
23	101010, 101100, 010000, 110110, 110011, 001000
32	111100, 011000, 010110, 110100, 101001, 000100.

In the second matrix, rows $3, 6$ determine source coordinates $2, 3$. Subtract row 2 from row 4 and remove the known coordinates to determine coordinate 5. Rows $1, 2, 5$ then determine coordinates $1, 4, 6$ in that order. In the third matrix row 6 determines coordinate 4. Row 4 then determines the sum of coordinates 1 and 2, and row 1 determines coordinate 3. Row 2 gives coordinate 2, and rows $3, 5$ give coordinates $5, 6$. The first matrix permutes coordinates. Hence all three maps are invertible. For the last block the formula is

$$L_{33}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h_0 & i \end{pmatrix} = \begin{pmatrix} b+f & c+e & c+e \\ a+c+h_0 & d+f+g & b+c+f \\ b+d+g+h_0 & a+e+h_0+i & f \end{pmatrix}.$$

For zero output, entries $(3,3), (1,1), (2,3), (1,2)$ give $f = b = c = e = 0$. Entry $(2,2)$ gives $d = g$, so entries $(3,1), (2,1), (3,2)$ give $h_0 = a = i = 0$. The kernel has dimension one, and the image is the eight-dimensional space of outputs A satisfying $A_{12} = A_{13}$. In even size use only the size-two pair to obtain an invertible map. In odd size use one size-three pair and size-two pairs on the remaining blocks. The image is the hyperplane $\ell_0(Au_0) = 0$, where u_0 is the sum of the second and third vectors in the size-three block and ℓ_0 is its first-coordinate functional. For any target T, choose a nonzero vector v and a nonzero functional ℓ killing v, Tv. This is possible since $n \geq 3$. Extend u_0 and v to bases of $\ker \ell_0$ and $\ker \ell$, respectively, and complete each basis by a vector of functional value one. The change-of-basis matrix P satisfies $Pu_0 = v$ and $\ell P = \ell_0$. Then $\ell_0(P^{-1}TPu_0) = \ell(Tv) = 0$. Conjugating inputs for $P^{-1}TP$ by P gives T.

For s_3, let E, P be square of size $n-1$, let C, G be rows and F a column of length $n-1$, and let $b \in K$. For the inputs below, the upper-left blocks of the six words sum to $3PE + 3EP + 2FC = [E, P]$. The upper-right blocks sum to $2PF + bF$ and the lower-left blocks to $2CE + 2GP + bG$. The lower-right block is CF, from the word xyz. Thus

$$s_3\left(\begin{pmatrix} P & 0 \\ C & b \end{pmatrix}, \begin{pmatrix} I_{n-1} & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} E & F \\ G & 0 \end{pmatrix}\right) = \begin{pmatrix} [E,P] & bF \\ bG & CF \end{pmatrix}.$$

For a nonscalar target W, choose v, Wv independent and a functional ℓ with $\ell(v) = 1$ and $\ell(Wv) = w = \operatorname{tr} W$. In the decomposition $\ker \ell \oplus Kv$, the target has the form $\left(\begin{smallmatrix} T & U \\ V & w \end{smallmatrix}\right)$ with $\operatorname{tr} T = 0$ and $U \neq 0$. Take $[E, P] = T$, $b = 1$, $F = U$, $G = V$ and choose C with $CU = w$. In even size, substituting I_n for the first input gives $s_3(I_n, R, S) = 3RS + 3SR = [R, S]$. Since $\operatorname{tr}(\lambda I_n) = 0$, Lemma 2.1 gives the scalar target λI_n. For odd n, take $[E, P] = \lambda I_{n-1}$, $b = 0$, $F = \lambda e_1$, $G = 0$, and $C = e_1^{\mathsf{T}}$, with e_1 the first coordinate vector of K^{n-1}. The formula gives λI_n. □

Proof of Theorem 1.3. The zero polynomial gives the zero image, and the size-one case follows from commutativity. For $n \geq 2$, the case $S \neq 0$ follows by setting two inputs to I_n. If $f \neq 0$ and $S = u = 0$, Proposition 2.2 and (3) give equality with $\mathfrak{sl}_n(K)$. When $S = 0$ and $u \neq 0$, Proposition 3.1 and Section 4 give every target over $|K| > 2$. Proposition 5.1 gives every target over $\mathbf{F}_2$. □

6 Matrix order two

We use the following elementary fact about multilinear maps into a plane [13, Lemma 3].

Lemma 6.1. *Let g be a multilinear map from a finite product of K-vector spaces to a two-dimensional K-vector space. If two values are linearly independent, g is surjective.*

Proof. Choose tuples with independent outputs that differ in as few positions as possible. A tuple obtained by exchanging some but not all of their differing entries is closer to both original tuples. Minimality forces its value to lie in the span of each original value, so it is zero. In every differing position take the sum of the two entries, except in one position where we take their linear combination with coefficients $\lambda, \mu \in K$. The expansion contains the original values with these coefficients, and every mixed term vanishes. If the tuples differ in one position, there are only the two original terms. □

Theorem 6.2. *For every multilinear quartic p over K, the image $p(M_2(K))$ is one of $\{0\}$, KI_2, $\mathfrak{sl}_2(K)$, and $M_2(K)$.*

Proof. Simultaneous conjugation of the inputs conjugates the value, and scaling one input scales the value by the same scalar. If all values are scalar, the image is zero or the scalar line. Suppose some value is nonscalar.

For a tuple of matrix units $E_{a_s b_s}$, $1 \le s \le 4$, any nonzero word product is a matrix unit E_{ij} with $j - i = \sum_{s=1}^{4}(b_s - a_s)$, by telescoping the indices. The integer $j - i$ is fixed for the tuple. In size two, $j - i = 0$ permits only diagonal units, while $j - i = 1$ or -1 permits only one off-diagonal unit. Thus each tuple gives a diagonal value or a multiple of one off-diagonal unit. If every matrix-unit evaluation were diagonal, multilinearity would make every value diagonal. For $a \neq b$, conjugating $\operatorname{diag}(a, b)$ by $I_2 + E_{12}$ gives $\operatorname{diag}(a, b) + (b - a)E_{12}$, ruling out that possibility for a polynomial with a nonscalar value. Thus scaling and swapping indices give E_{12} and E_{21} in the image.

The diagonal and off-diagonal subspaces give a grading by $\mathbb{Z}/2\mathbb{Z}$, with the diagonal subspace of degree zero. Restrict each input to the subspace containing its matrix unit in a tuple giving E_{12}. The sum of the four degrees is one, so every output of the restriction is off diagonal. Swapping indices preserves all four input subspaces and gives E_{21} in the same restriction. Lemma 6.1 yields the whole off-diagonal subspace. Every nonscalar traceless $T \in M_2(K)$ is similar to an off-diagonal matrix, using the basis v, Tv from Lemma 2.1.

Characteristic two. Suppose first that p has only traceless values. Conjugating E_{12} by $I_2 + E_{21}$ gives $E_{12} + E_{21} + I_2$. Expand the inputs for this value in matrix units. Since its diagonal part is nonzero, a tuple of matrix units gives a nonzero diagonal value. This value is traceless, hence scalar in characteristic two. Scaling gives all scalars and finishes the traceless case.

If p has a value of nonzero trace, multilinearity gives a matrix-unit tuple with value $\operatorname{diag}(a, b)$, $a + b \neq 0$. Restrict each input to the diagonal or off-diagonal subspace according to this tuple. The image of the restriction is diagonal and contains $\operatorname{diag}(b, a)$ by simultaneous index exchange. The two values are independent since $a^2 - b^2 = (a + b)^2 \neq 0$. Lemma 6.1 gives all diagonal matrices.

For each separable quadratic extension L/K, let σ be its nontrivial K-automorphism and let $m_a : L \to L$ send x to ax for $a \in L$. Identify $M_2(K)$ with $\operatorname{End}_K(L)$ in a basis of L. The subspaces of maps m_a and $m_b\sigma$, $a, b \in L$, give another grading by $\mathbb{Z}/2\mathbb{Z}$. Indeed, write $L = K(\theta)$ with $\theta^2 + \theta = c \in K$ and $\sigma(\theta) = \theta + 1$. In the basis $1, \theta$, the two maps associated to $x + y\theta$ are

$$\begin{pmatrix} x & cy \\ y & x + y \end{pmatrix}, \qquad \begin{pmatrix} x & x + cy \\ y & x \end{pmatrix}.$$

Comparison of the first columns and the lower-right entries shows that an element in the intersection has $y = 0$. Comparison of the upper-right entries then gives $x = 0$. The two subspaces are thus complementary, and $\sigma m_a = m_{\sigma(a)}\sigma$ gives the grading. The first map has trace $a + \sigma(a)$ for $a = x + y\theta$, and the second map has trace zero.

Expand a nonzero-trace value of p by decomposing each input into its two homogeneous components. A homogeneous tuple gives a value m_a with $a+\sigma(a) \neq 0$. On this tuple's fixed pattern of input degrees, the outputs all lie in the two-dimensional subspace of multiplication maps. Conjugation by σ preserves each input subspace and supplies $m_{\sigma(a)}$ in the same restriction. These two values are independent. Indeed, if $\sigma(a) = c_0 a$ for $c_0 \in K$, then $c_0^2 = 1$, hence $c_0 = 1$, contrary to $a + \sigma(a) \neq 0$. Lemma 6.1 gives every multiplication map.

A matrix of nonzero trace has separable characteristic polynomial in characteristic two. If it splits, the matrix is diagonalizable over K and occurs by the diagonal restriction above. Otherwise its characteristic polynomial is irreducible, and its companion form is a multiplication map on the corresponding quadratic field. All traceless targets are already covered by the diagonal and off-diagonal restrictions.

Characteristic different from two. The only scalar traceless matrix is zero. Thus if p has only traceless values, the first part of the proof gives image $\mathfrak{sl}_2(K)$. Suppose now that p has some value of nonzero trace.

Consider bases I_2, H, J, HJ of $M_2(K)$ with $H^2 = aI_2$, $J^2 = bI_2$, $HJ = -JH$, where $a, b \in K^\times$. One such basis has $H = \operatorname{diag}(1, -1)$ and $J = E_{12} + E_{21}$. Assign their degrees $0, h, j, h + j$ in $(\mathbb{Z}/2\mathbb{Z})^2$, with generators h, j, respectively. For a tuple $\mathbf{g} = (g_1, \ldots, g_4)$ of degrees, let U_{g_i} be the basis element of degree g_i. The value of p is

$$p(U_{g_1}, \ldots, U_{g_4}) = \lambda(\mathbf{g}) U_{g_1} U_{g_2} U_{g_3} U_{g_4}.$$

Here $\lambda(\mathbf{g})$ is the sum of the word coefficients with the signs obtained by reordering each word into $U_{g_1} U_{g_2} U_{g_3} U_{g_4}$. Distinct nonzero degrees anticommute, and all other pairs commute, so $\lambda(\mathbf{g})$ is independent of a, b. In particular, the tuple gives a nonzero value precisely when $\lambda(\mathbf{g}) \neq 0$, because each basis input is invertible. If nonzero, its value is scalar when the total degree is zero and lies in one of the three traceless basis directions otherwise. The same statements hold after scaling individual basis inputs by nonzero scalars. They also hold under any permutation of the three nonzero degree names, since their commutation signs are preserved.

Expanding inputs in this basis, we find a tuple giving a nonzero scalar value and another giving a nonscalar value. Choose a pair of degree tuples of these two types with the smallest number r of differing positions. Name the output degree of the nonscalar tuple h. In any of the displayed bases, let $\mathbf{A} = (A_1, \ldots, A_4)$ and $\mathbf{B} = (B_1, \ldots, B_4)$ be the corresponding input tuples, and let $S \subseteq \{1, 2, 3, 4\}$ be the set of differing positions. Write

$$p(\mathbf{A}) = \alpha I_2, \qquad p(\mathbf{B}) = \beta H, \qquad \alpha\beta \neq 0.$$

The set of degree tuples with nonzero evaluation and the minimum $r = |S| \geq 1$ are independent of the basis parameters. Every proper mixture of the endpoints in the positions S has zero value. Otherwise, depending on whether it is scalar or nonscalar, it would be closer to an endpoint of the opposite type. Use $A_s + B_s$ in every position $s \in S$ except one, where we use $\lambda A_s + \mu B_s$, with $\lambda, \mu \in K$. The only terms of the expansion with possibly nonzero values are the two original tuples. Their contribution $\lambda\alpha I_2 + \mu\beta H$ ranges over $KI_2 + KH$. For any $a \neq 0$, we can take $H = \left(\begin{smallmatrix} 0 & a \\ 1 & 0 \end{smallmatrix}\right)$ and $J = \operatorname{diag}(1, -1)$. For a nonscalar target T, the traceless part $T - \frac{1}{2}(\operatorname{tr} T) I_2$ has square aI_2 for some $a \in K$. When $a \neq 0$, it is similar to the indicated H by using a basis $v, (T - \frac{1}{2}(\operatorname{tr} T) I_2)v$. Thus the planes give all such targets and all scalars.

It remains to produce a value $\alpha I_2 + N$ with $\alpha \neq 0$ and N nonzero nilpotent. By applying the commutator rule $[J, UV] = [J, U]V + U[J, V]$ to each word,

$$[J, p(\mathbf{B})] = \sum_{\ell=1}^{4} p(B_1, \ldots, B_{\ell-1}, [J, B_\ell], B_{\ell+1}, \ldots, B_4).$$

The left side is $2\beta JH \neq 0$. Choose a position ℓ giving a nonzero summand. The inserted commutator is $2JB_\ell$, and moving J to the left of the ordered product introduces only a sign. Thus the summand equals $\delta\beta JH$ for $\delta \in K^\times$, where δ is ± 2 times the ratio of the ordered-product coefficients of the modified and original degree tuples. The nonzero summand and δ are independent of a, b. In what follows, $\eta \in K$ is the coefficient of the commutator input. Without it, we use $A_s + B_s$ in each differing position $s \in S$ and A_s elsewhere, giving $\alpha I_2 + \beta H$.

If $\ell \in S$, add $\eta[J, B_\ell]$ to the ℓ-th input. The degree of $[J, B_\ell]$ differs from both original degrees in that position. It differs from that of B_ℓ by $j \neq 0$. Equality with the degree of A_ℓ would give a nonscalar modified endpoint at distance $r - 1$ from the scalar endpoint. In the coefficient of η, a term omitting any of the other choices from $\mathbf{B}$ has distance less than r from $\mathbf{A}$ and from the modified nonscalar endpoint. Hence it cannot have a nonzero value of either type. The result is the value $\alpha I_2 + \beta(H + \eta\delta JH)$.

If $\ell \notin S$, add $\eta[J, A_\ell]$ to the common ℓ-th input. In the coefficient of η, consider a term with t choices from $\mathbf{B}$ among the r positions in S. A scalar term with $t \geq 1$ must be zero, since its distance to the modified nonscalar endpoint is $r - t < r$. For $t = 0$ the degree is j, so this term also cannot be a nonzero scalar. Every remaining nonzero term is nonscalar and has distance $t + 1$ from $\mathbf{A}$, so $t \geq r - 1$. The full choice has value $\delta\beta JH$. Write deg for the degree of a nonzero homogeneous element in this grading. A single omission at position $s \in S$ has degree $h + j + g_s$, where $g_s = \deg A_s + \deg B_s \neq 0$. If $g_s = h + j$ its value is scalar and vanishes. If $g_s = j$ or $g_s = h$, the term lies in KH or KJ, respectively.

For each $s \in S$, take the representative $B_s = A_s U_{g_s}$ of its degree. This changes B_s by a nonzero scalar and preserves the vanishing of all mixed evaluations used above. Use β for the resulting coefficient of H in $p(\mathbf{B})$. The commutator ratio δ is unaffected. The ordered product in the omitted term replaces B_s by $B_s U_{g_s}^{-1}$ and B_ℓ by $2JB_\ell$. Moving J and $U_{g_s}^{-1}$ to the left gives $\pm 2JU_{g_s}^{-1}B_1B_2B_3B_4$. The coefficient of the modified ordered product depends only on the new tuple of degrees. Dividing by the coefficient of the original $\mathbf{B}$ tuple, this term is therefore a scalar multiple of $\beta JU_{g_s}^{-1}H$ with multiplier independent of a, b. For $g_s = j$ we have $JJ^{-1}H = H$, and for $g_s = h$ we have $JH^{-1}H = J$. Summing the coefficients for single omissions gives constants $\rho, \varepsilon \in K$ independent of a, b. Thus the value in this case is

$$\alpha I_2 + \beta\big((1 + \eta\rho)H + \eta\varepsilon J + \eta\delta JH\big).$$

The same formula holds in the differing-position case with $\rho = \varepsilon = 0$.

Choose $\eta \neq 0$ with $1 + \eta\rho \neq 0$, and $a \neq 0$ with $\varepsilon^2 - a\delta^2 \neq 0$. Each choice excludes at most one element of $K^\times$, which has at least two elements. Take $H = \left(\begin{smallmatrix} 0 & a \\ 1 & 0 \end{smallmatrix}\right)$, put $D = \operatorname{diag}(1, -1)$, and set

$$J = \frac{1 + \eta\rho}{\eta} H(\varepsilon I_2 - \delta H)^{-1} D.$$

The inverse exists because $(\varepsilon I_2 - \delta H)(\varepsilon I_2 + \delta H) = (\varepsilon^2 - a\delta^2) I_2$. The relation $DH = -HD$ gives $JH = -HJ$ and

$$J^2 = -\frac{a(1 + \eta\rho)^2}{\eta^2(\varepsilon^2 - a\delta^2)} I_2.$$

The four matrices I_2, H, D, HD form a basis by comparison of their diagonal and off-diagonal entries. Multiplying D, HD on the left by the invertible matrix $\frac{1+\eta\rho}{\eta} H(\varepsilon I_2 - \delta H)^{-1}$ in the span of I_2, H replaces them by J, HJ and preserves this basis property. Also $J(\varepsilon I_2 +$

$\delta H) = \frac{1+\eta\rho}{\eta} HD$, so the difference from αI_2 in the constructed value is $\beta(1+\eta\rho)H(I_2+D)$. This is nonzero since H is invertible and $I_2 + D = \mathrm{diag}(2,0)$. Its square is zero since $H(I_2+D)H(I_2+D) = a(I_2-D)(I_2+D) = 0$.

In the resulting value $\alpha I_2 + N$, both α and N are nonzero. Scaling prescribes the nonzero scalar part. Every nonzero square-zero matrix of size two is similar to E_{12}, using the basis Nv, v with $Nv \neq 0$. Hence conjugation prescribes the nonzero nilpotent part, completing the image $M_2(K)$. □

7 Block sources for quartics

For a fixed polynomial, take block diagonal direct sums in each input position. Each word evaluates to the direct sum of its evaluations on the original tuples. We use this property with the cyclic decomposition of a matrix over K. In a cyclic basis of a summand of dimension m, the matrix of T is the column companion of its characteristic polynomial. Its first $m-1$ columns are successive coordinate vectors $e_2, \ldots, e_m$, and its last column lists the negatives of the coefficients in ascending degree below m. Rational canonical form decomposes an arbitrary operator into cyclic summands whose characteristic polynomials form a divisibility chain.

Lemma 7.1. *Every operator T in even dimension is a direct sum of scalar blocks of order two and even-dimensional operators admitting a subspace V with $V \oplus T(V)$ equal to the whole summand.*

Proof. Take the cyclic summands of a target in rational canonical form. Retain each even cyclic summand and pair the odd cyclic summands, whose number is even. On an even cyclic summand, the span V of the cyclic basis vectors with even exponents satisfies $V \oplus T(V)$ equal to the whole summand.

For a pair of odd cyclic summands, write their cyclic bases as $v_0, \ldots, v_{2r}$ and $w_0, \ldots, w_{2s}$. If $r = s = 0$, the pair is either scalar, or $V = K(v_0 + w_0)$ gives the desired decomposition. Otherwise interchange the summands if needed so that $r \geq 1$. Let V_0 be the span of the even-indexed vectors in the two lists excluding v_{2r} and w_{2s}. The space $V_0 \oplus T(V_0)$ is spanned by all basis vectors except those two. Let α and β be the coefficients of v_{2r} and w_{2s} in Tv_{2r} and Tw_{2s}, respectively. Set

$$v = v_{2r} + w_{2s} + cv_{2r-1}, \qquad V = V_0 + Kv,$$

where $c \in K$ is chosen with $\alpha + c \neq \beta$. In the quotient by $V_0 \oplus T(V_0)$, the coordinate pairs of v, Tv are $(1,1)$ and $(\alpha + c, \beta)$. They are independent, so $V \oplus T(V)$ is the whole pair of summands. □

Lemma 7.2. *Let p be an ordinary multilinear quartic over K. Suppose that for fixed $a, b \in K$ with $a + b \neq 0$, for each $h \geq 1$ the image of p on $M_{2h}(K)$ contains the matrices*

$$\begin{pmatrix} aA & I_h \\ B & bA \end{pmatrix}, \qquad A, B \in M_h(K),$$

and that the image on $M_2(K)$ contains all scalars. Then p is surjective in every even matrix order.

Proof. Decompose a target by Lemma 7.1. On each piece with a half-dimensional subspace V as in the lemma, choose a basis of V followed by its T-images. The matrix of the restriction

has the form $T_0 = \left(\begin{smallmatrix} 0 & B_0 \\ I_h & D \end{smallmatrix}\right)$, with $B_0, D \in M_h(K)$. Put $A = D/(a+b)$ and $S = \left(\begin{smallmatrix} I_h & -bA \\ 0 & I_h \end{smallmatrix}\right)$. Then

$$S^{-1}T_0S = \begin{pmatrix} bA & B_0 + abA^2 \\ I_h & aA \end{pmatrix}.$$

Interchanging the two blocks gives a matrix in the hypothesized family. A scalar pair is available by hypothesis. Similarity preserves polynomial images, and taking direct sums of the input tuples gives the whole target. □

Lemma 7.3. *Let p be an ordinary multilinear quartic over K. For $h \geq 1$ write*

$$J_h = \begin{pmatrix} I_h \\ 0 \end{pmatrix}, \qquad C_h = \begin{pmatrix} 0 \\ I_h \end{pmatrix} \quad \text{in } M_{(h+1)\times h}(K).$$

Suppose that for some $a, b \in K$ with $a(a+b) \neq 0$ and every $h \geq 1$, the image on $M_{2h+1}(K)$ contains either the entire family

$$\begin{pmatrix} aJ_h[A & w] & C_h \\ B & & bA \end{pmatrix}, \quad A \in M_h(K), \quad w \in K^h, \quad B \in M_{h\times(h+1)}(K), \tag{11}$$

or the transposes of the entire family. Here $[A\ \ w]$ appends the column w to A. Alternatively, for fixed $a, d \in K$ with $a \neq 0$, we allow the same family with the bottom-right block replaced by dI_h, or its transposes. Then this image contains every odd cyclic class of order at least three, and every direct sum of a scalar singleton with a cyclic block of even order at least four. If p is surjective in every even order and in order three, it is surjective in every order at least two.

Proof. A matrix is similar to its transpose over K by rational canonical form, so either hypothesis suffices. Write $e_1, \ldots, e_{h+1}$ for the coordinate vectors of the first block, and $f_1, \ldots, f_h$ for those of the second. Interleave them in the order $e_1, f_1, \ldots, e_h, f_h, e_{h+1}$. Take $A = \mathrm{diag}(z, 0, \ldots, 0)$, $B_{i,i} = 1$ for $1 \leq i \leq h$, $B_{1,j} = u_j$ for $2 \leq j \leq h+1$, and the remaining entries of B zero. The matrix in this basis has all subdiagonal entries one and all entries below the subdiagonal zero. Its other possibly nonzero entries are

$$M_{11} = az, \quad M_{22} = bz, \quad M_{2,2j-1} = u_j \ (2 \leq j \leq h+1), \quad M_{2k-1,\,2h+1} = aw_k \ (1 \leq k \leq h),$$

where w_k is the k-th coordinate of w. Such an upper Hessenberg matrix of order m is cyclic, since $M^i e_1$ has entry one at position $i+1$ and zeros below it for $0 \leq i < m$ in the interleaved basis.

For an entry x in position (i, j) with $i < j$, its first contribution to $\det(tI_m - M)$, in decreasing powers of t, is $-xt^{m-j+i-1}$. In the determinant expansion, a directed permutation cycle using the edge $j \to i$ must return from i to j along edges that increase the index by at most one. The unique shortest such return uses the subdiagonal edges $i \to i+1 \to \cdots \to j$. The cycle of length $j-i+1$ contributes the permutation sign $(-1)^{j-i}$ and the entry product $(-1)^{j-i+1}x$. The highest-degree term fixes every other vertex and chooses t at each of them. Every other term using x has smaller t-degree. The coefficient of t^{m-1} for our matrix is $-(a+b)z$. In the fixed-block alternative, the only change is that the diagonal entries in the f_i coordinates equal d, so this coefficient is $-az - hd$. The first degrees for u_j and w_k are $m - 2j + 2$ and $2k - 2$, with coefficients -1 and $-a$, respectively. For $m = 2h+1$ these give every degree from $m-2$ to zero exactly once. The coefficient of z is nonzero in either case. Set z from the coefficient of t^{m-1}, then solve for the remaining variables in decreasing order of their first degrees. At the first degree of each variable it occurs only through the corresponding shortest cycle, and variables with smaller first degrees are absent.

The coefficient at that degree is therefore a nonzero linear function of the variable plus an already fixed term. This produces every cyclic class of order m.

For a singleton with scalar λ and a cyclic block of order $2h$, $h \geq 2$, instead set $A = \operatorname{diag}(z_1, z_2, 0, \ldots, 0)$ with $z_1 = \lambda/a$, and change B_{11} and w_1 to zero. The first coordinate of the interleaved matrix is then a direct summand with value λ. The remaining matrix is upper Hessenberg with subdiagonal entries one and trace $bz_1 + (a+b)z_2$, or $az_2 + hd$ in the fixed-block alternative. For u_j and w_k with $k \geq 2$, the first characteristic degrees are $2h-2j+2$ and $2k-3$ with coefficients -1 and $-a$. They again cover the degrees from $2h-2$ to zero once each, so the remaining block attains every cyclic class.

Assume now the additional surjectivity in the stated orders and take a target of odd order at least three. If some cyclic summand has odd order at least three, use the first construction for that summand and surjectivity in even order for the complement, if nonempty. Otherwise the odd cyclic summands are singletons. If there is an even cyclic summand of order at least four, combine it with a singleton and use the second construction, again with an even complement. In the remaining case, take either three singletons or a singleton and a block of order two as the order-three summand, leaving an even complement. Direct sums of the original inputs give the target in every case. □

8 Proper quartics

Use left-normed brackets for the commutator $UV - VU$, and write an input index i in a bracket for X_i. For example, $[1,2,3,4] = [[[X_1, X_2], X_3], X_4]$. Define

$$\begin{aligned} L_1 &= [1,2,3,4], \quad L_2 = [1,2,4,3], \quad L_3 = [1,3,2,4], \\ L_4 &= [1,3,4,2], \quad L_5 = [1,4,2,3], \quad L_6 = [1,4,3,2], \\ P_1 &= [1,2][3,4], \quad P_2 = [1,3][2,4], \quad P_3 = [1,4][2,3]. \end{aligned}$$

Lemma 8.1. *The nine displayed polynomials are a basis of the proper quartics over every field. Write a proper quartic in this basis as $p = \sum_{i=1}^{6} l_i L_i + t_1 P_1 + t_2 P_2 + t_3 P_3$. When all but the i-th input are set to a common matrix A, the resulting value with X in position i is $q_i[A,[A,[A,X]]]$, where*

$$(q_1, q_2, q_3, q_4) = \big(-\textstyle\sum_{j=1}^{6} l_j,\ l_1 + l_2,\ l_3 + l_4,\ l_5 + l_6\big).$$

All four q_i vanish exactly when p belongs to the span of the six ordered products

$$P_1,\ [3,4][1,2],\ P_2,\ [2,4][1,3],\ P_3,\ [2,3][1,4].$$

For every $n \geq 3$, the trace of p vanishes on $M_n(K)$ exactly when (t_1, t_2, t_3) is proportional to $(1, -1, 1)$.

Proof. The four deletion maps on the 24 free words have matrix entries 0 and 1. A row (i, w) records the coefficient of the three-letter word w after deleting i. On the rows

$$\begin{gathered} (1,234), (1,243), (1,324), (1,342), (1,423), (1,432), \\ (2,134), (2,143), (2,314), (2,413), (3,124), \\ (3,142), (3,412), (4,123), (4,132) \end{gathered}$$

and columns

$$1234, 1243, 1324, 1342, 1423, 1432, 2134, 2314, 2341, 2413, 2431, 3124, 3214, 3412, 3421.$$

The entry in row (i, w) and column v is 1 if deleting i in v gives w, and 0 otherwise. For the determinant, eliminate below each diagonal entry from left to right, moving the first available nonzero entry in the column to the pivot position. The row exchanges are

$$(7, 11),\ (8, 9),\ (9, 11),\ (12, 15),\ (13, 14),$$

in the current row order. The pivots, before scaling, are

$$1, 1, 1, 1, 1, 1, -1, 1, -1, 1, -1, -1, 1, 1, -1.$$

There are five exchanges and five negative pivots, so the determinant is 1. The proper space has dimension at most nine. Each displayed bracket polynomial is proper. For the coefficients of the nine proposed basis elements, use the rows

$$1234, 1243, 1324, 1342, 1423, 1432, 2314, 2413, 3412.$$

The first six rows are the identity on the L_i columns and have the additional entries $1, -1$ in the successive P_j columns, and the last three rows have ones in L-columns $3, 4, 6$, $4, 5, 6$, and $2, 5, 6$, respectively, with all other entries zero. Subtracting their indicated sums of the first six rows makes the last three L-rows zero and leaves the last three P-rows $(0, 0, 1)$, $(0, 1, 0)$, and $(1, 0, 0)$. Their determinant is -1. This proves the basis assertion in all characteristics.

If the single distinct input is in position 1, every L_j becomes

$$[[[X, A], A], A] = -[A, [A, [A, X]]].$$

If the distinct input is second in the left-normed bracket, its sign is positive. If it is third or fourth, the first commutator vanishes. Every P_j contains a commutator of two equal inputs. These give the four q_i. If all $q_i = 0$, the Lie part is a combination of $L_1 - L_2, L_3 - L_4, L_5 - L_6$. By $[[U, V], W] - [[U, W], V] = [U, [V, W]]$, these three differences are $[[1, 2], [3, 4]]$, $[[1, 3], [2, 4]]$, and $[[1, 4], [2, 3]]$, respectively. Each is the difference of two of the six ordered products. Every ordered product vanishes in each of the four one-variable restrictions.

The trace of every L_i is zero. By $\operatorname{tr}([A, B]C) = \operatorname{tr}(A[B, C])$,

$$\operatorname{tr}(P_1 - P_2 + P_3) = \operatorname{tr}\big(X_1([X_2, [X_3, X_4]] - [X_3, [X_2, X_4]] + [X_4, [X_2, X_3]])\big) = 0$$

by Jacobi. For the two tuples $(E_{12}, E_{21}, E_{13}, E_{31})$ and $(E_{12}, E_{13}, E_{21}, E_{31})$, the three products are respectively $(E_{11}, 0, -E_{33})$ and $(0, E_{11}, E_{33})$, by $E_{ij}E_{k\ell} = \delta_{jk}E_{i\ell}$, where $\delta_{jk} = 1$ for $j = k$ and zero otherwise. Taking their traces gives $t_1 - t_3 = 0$ and $t_2 + t_3 = 0$ for a trace-zero quartic. □

8.1 A proper quartic with a nonzero one-variable coefficient

Fix two input positions, with matrices X, Y. Set the other two inputs to $P = \operatorname{diag}(I_u, 0_v)$ for block sizes $u, v \geq 1$. Write X_{ij}, Y_{ij}, F_{ij} for the blocks of X, Y and the value F, with $i, j \in \{1, 2\}$. A word evaluated in blocks is a sum of products in which consecutive column and row blocks agree. In the first diagonal block, a product that leaves the first summand must return to it using $X_{12}Y_{21}$ or $Y_{12}X_{21}$. Its coefficient vanishes if either copy of P lies between X, Y, and otherwise is its word coefficient. Let r, s be the sums of the surviving coefficients in the two orders, respectively. In the first summand the terms that stay in that summand sum to p with the two fixed inputs equal to I_u, which is zero by properness. In the second summand those terms are zero because the two fixed blocks are zero. Denote the one-variable coefficients of the selected positions by q_X, q_Y.

Lemma 8.2. *With these notations, the four output blocks are*

$$\begin{aligned}
F_{11} &= rX_{12}Y_{21} + sY_{12}X_{21},\\
F_{22} &= rX_{21}Y_{12} + sY_{21}X_{12},\\
F_{12} &= q_Y(X_{11}Y_{12} - Y_{12}X_{22}) + q_X(Y_{11}X_{12} - X_{12}Y_{22}),\\
F_{21} &= q_Y(X_{22}Y_{21} - Y_{21}X_{11}) + q_X(Y_{22}X_{21} - X_{21}Y_{11}).
\end{aligned}$$

For a proper quartic with nonzero formal trace and some $q_i \neq 0$, the two positions can be chosen with $q_X \neq 0$ and $r + s \neq 0$.

Proof. In the second diagonal block the two return coefficients are again r, s, because replacing P by $I-P$ in both fixed slots leaves the value unchanged by multilinearity and properness. The coefficient of $X_{11}Y_{12}$ in F_{12} is q_Y by taking $X = P$ and using $[P,[P,[P,Y]]]_{12} = Y_{12}$. Setting $X = I$ forces the coefficient of $Y_{12}X_{22}$ to be its negative. The same reasoning with $Y = P, I$ gives the other two terms in F_{12}. Exchanging the two summands gives F_{21}.

For positions i, j, $r+s$ is the trace obtained by putting E_{12}, E_{21} there and E_{11} elsewhere in order two. The L_i have zero trace. The two elementary commutators $[E_{12}, E_{11}] = -E_{12}$, $[E_{21}, E_{11}] = E_{21}$ show that the paired products P_1, P_2, P_3 have traces $(0, -1, -1)$ for pairs $12, 34$, $(-1, 0, 1)$ for $13, 24$, and $(1, 1, 0)$ for $14, 23$. In each triple, the zero occurs where two copies of E_{11} are in the same commutator. Hence for any fixed position the three trace sums are

$$-t_2 - t_3, \qquad -t_1 + t_3, \qquad t_1 + t_2$$

in some order. They cannot all be zero for a traceful proper quartic by Lemma 8.1. Choose a position with $q_i \neq 0$ and a partner giving nonzero return sum. □

Proposition 8.3. *A proper quartic with nonzero formal trace and some $q_i \neq 0$ is surjective in every order at least two.*

Proof. Choose positions as in Lemma 8.2 and put $q = q_X$. In equal blocks of order h, take

$$X = \begin{pmatrix} 0 & I_h \\ -B & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} q^{-1}I_h & 0 \\ A & 0 \end{pmatrix}.$$

The output is $\left(\begin{smallmatrix} rA & I_h \\ B & sA \end{smallmatrix}\right)$. With $X_{12} = X_{21} = I_h$ and $Y_{12} = Y_{21} = \lambda I_h/(r+s)$, all diagonal blocks of the inputs zero, the output is λI_{2h}. This proves surjectivity in even order by Lemma 7.2.

For block orders $h+1, h$, use J_h, C_h from Lemma 7.3. Let R be the cyclic permutation on the first block with $RJ_h = C_h$. When $r \neq 0$, put $X_{11} = X_{22} = Y_{22} = Y_{12} = 0$, and

$$X_{12} = J_h, \quad Y_{11} = q^{-1}R, \quad Y_{21} = [A \;\; w], \quad X_{21} = -q^{-1}BY_{11}^{-1}.$$

The output is the odd family with $a = r, b = s$. When $r = 0$, use the transposes of these four blocks in the opposite block positions, with $Y_{11} = -q^{-1}R^{\mathsf{T}}$, $X_{12} = q^{-1}Y_{11}^{-1}B^{\mathsf{T}}$, and keep the other blocks zero. That is, $X_{21} = J_h^{\mathsf{T}}$ and $Y_{12} = [A \;\; w]^{\mathsf{T}}$. The output is the transpose of the odd family with $a = s, b = 0$. Lemma 7.3 now leaves only the order-three noncyclic targets.

A noncyclic matrix of order three is either scalar or a scalar matrix λI_3 plus a matrix T of rank one. Indeed its invariant factors have degrees one and two with the former dividing the latter, or all have degree one. Writing the rank-one matrix as a column times a row gives $T^2 = (\operatorname{tr} T)T$. If $\operatorname{tr} T \neq 0$, its image is the eigenspace of eigenvalue $\operatorname{tr} T$ and its kernel is the zero eigenspace. If $\operatorname{tr} T = 0$, it has one nilpotent block of size two and one zero singleton. Thus the class of $\lambda I_3 + T$ is determined by λ and $\mu = \lambda + \operatorname{tr} T$.

If K has at least three elements, choose distinct $a_0, a_1, a_2 \in K$. Take both auxiliary matrices equal to $\mathrm{diag}(a_0, a_1, a_2)$. With indices modulo three, put $X_{i,i+1} = 1$ and $Y_{i+1,i} = w_i$, all other entries zero. On a return path between vertices i, j, the two auxiliary positions contribute a coefficient $r(a_i - a_j)^2$ for order XY and $s(a_i - a_j)^2$ for order YX. To see this, subtract $a_j I$ from both auxiliary inputs using properness. Deleting an auxiliary letter preserves the order of X, Y, so this subtraction also applies to each ordered contribution. In the word sum the two auxiliary entries are now zero at vertex j and $a_i - a_j$ at vertex i, so the surviving terms are exactly those for the projection coefficient multiplied by $(a_i - a_j)^2$. The output is diagonal with entries

$$r(a_i - a_{i+1})^2 w_i + s(a_{i-1} - a_i)^2 w_{i-1}.$$

Setting $w_i = \lambda/((r+s)(a_i - a_{i+1})^2)$ gives λI_3.

For $K = \mathbf{F}_2$, let $L = \mathbf{F}_8$, $\sigma(z) = z^2$, and choose $\alpha \in L \setminus K$. Write m_z for multiplication by z on L over K. Put both auxiliary inputs equal to m_α, $X = \sigma$, and $Y = m_u \sigma^{-1}$. Moving multiplication operators past σ by $\sigma m_z = m_{\sigma(z)}\sigma$, each word with X before Y becomes multiplication by a product with α in the two outer regions of the word, $\sigma(\alpha)$ in the middle region, and a factor $\sigma(u)$. For Y before X, the corresponding scalars are $\alpha, \sigma^{-1}(\alpha), \alpha$ and u. Applying the same return coefficient formula therefore gives

$$m_{r(\alpha-\sigma(\alpha))^2\sigma(u)+s(\alpha-\sigma^{-1}(\alpha))^2 u}.$$

The coefficients (r, s) are either $(1, 0)$ or $(0, 1)$, so a choice of u gives I_3. Then scaling one input gives every scalar value.

For the rank-one targets, if $\lambda = 0$ or $\mu = 0$, there is a zero singleton summand and an order-two complement, covered by even-order surjectivity and a zero singleton source. Suppose now $\lambda\mu \neq 0$. With block orders $2, 1$, set $X_{11} = X_{22} = Y_{22} = 0$ and write

$$X_{12} = u, \quad X_{21} = v, \quad Y_{12} = x, \quad Y_{21} = y, \quad Y_{11} = q^{-1}D.$$

Here u, x are columns of length two, v, y are rows of length two, and D has order two. The output is

$$F = \begin{pmatrix} ruy + sxv & Du \\ -vD & rvx + syu \end{pmatrix}.$$

If $r^2 \neq s^2$, take

$$u = (1, \lambda)^{\mathsf{T}}, \quad v = (0, 1), \quad D = \begin{pmatrix} 1 & 0 \\ \mu\lambda & -\mu \end{pmatrix}, \quad y = cv, \quad x = du,$$

where $rc + sd = 1$, $sc + rd = (\lambda + \mu)/\lambda$. These equations are solvable since $r^2 \neq s^2$. The output satisfies $F - \lambda I_3 = (1, 0, \mu)^{\mathsf{T}}(-\lambda, 1, 1)$, which has rank one and trace $\mu - \lambda$. Otherwise $r = s \neq 0$ and the characteristic differs from two. Put $\kappa = \mu/2$, $\tau = -1/\kappa^2$, and take

$$u = \tau(1, \lambda)^{\mathsf{T}}, \quad v = (1, 0), \quad x = (\kappa/s, 0)^{\mathsf{T}}, \quad y = (0, 1/(r\tau)), \quad D = \begin{pmatrix} -\kappa(\kappa - \lambda) & -\kappa \\ 0 & 0 \end{pmatrix}.$$

Then $F - \lambda I_3 = (1, 0, \kappa)^{\mathsf{T}}(\kappa - \lambda, 1, 1)$, again of rank one and trace $\mu - \lambda$. This completes order three, and Lemma 7.3 completes the proof. □

Proposition 8.4. *If a proper quartic is trace-zero and some $q_i \neq 0$, its image in every order at least three is the traceless subspace.*

Proof. By Lemma 2.1, every nonscalar traceless target is $[A, B]$ for some regular semisimple A. Over a splitting field, the eigenvalues of $\mathrm{ad}_A(X) = [A, X]$ on matrix units are the differences of the eigenvalues of A. Thus its image and the image of its third power have equal dimension. These ranks agree over K, and containment of the latter image in the former gives equality. Thus the one-variable restriction $q_i \,\mathrm{ad}_A^3$ attains the target.

A nonzero scalar matrix λI_n has trace zero exactly when the characteristic is a prime dividing n. In characteristic two, putting all but the i-th input equal to E_{11} and the remaining one to E_{12} gives $q_i E_{12}$ in order two. The trace vanishes also in size two because each L_i is a commutator and the trace of $P_1 - P_2 + P_3$ vanishes in every order. Theorem 6.2 gives I_2 in its image, and block sums and scaling give the required scalar targets.

In characteristic $\ell \geq 5$, take the order-ℓ Weyl pair A, D from the cubic proof, so $[A, D] = I_\ell$. The derivation property gives $\mathrm{ad}_A(D^j) = jD^{j-1}$ for $j \geq 1$ by expanding the commutator of A with a product and using $[A, D] = I_\ell$. Applying it three times to D^3 gives $6I_\ell$. Scale D^3 in the free input and take block sums.

In characteristic three, the proper trace-zero space has basis $L_1, \ldots, L_6, R$ with $R = P_1 - P_2 + P_3$. Put $J = E_{12} + E_{23}$ and $V = E_{21} + 2E_{32}$ in order three. The table gives the coefficients of I_3 obtained at the displayed tuples. To compute a column, expand each bracket by $UV - VU$, use $E_{ij}E_{k\ell} = \delta_{jk}E_{i\ell}$ in each product, and add modulo three. For example, in the first row $L_5 = [[[V, J^2V], V], J]$. Since $[J, V] = I_3$, its successive brackets are $-2JV$, $-2V$, and $2I_3$.

Inputs	L_1	L_2	L_3	L_4	L_5	L_6	R
(V, V, J, J^2V)	0	0	0	0	2	2	0
(V, V, JV, J^2)	0	0	0	2	0	1	2
(V, V^2, J, J^2)	0	0	0	0	1	0	2
(V, V, J^2V, J)	0	0	2	2	0	0	0
(V, JV, V, J^2)	0	2	0	0	1	0	1
(V, JV, JV, J)	2	0	2	0	0	0	0

Eliminating below the diagonal in the first six columns over $\mathbf{F}_3$ gives the row exchanges and pivots

$$(1, 6),\ (2, 5),\ (3, 4),\ (4, 5), \qquad 2, 2, 2, 2, 1, 2.$$

Thus their determinant is 2. Multiplication shows that the vector $(1, 2, 2, 1, 1, 2, 1)$ spans the kernel. All four one-variable coefficients vanish on this line by Lemma 8.1, so some row gives a nonzero scalar value for p. Scale one input and take order-three block sums to obtain the target. Trace vanishing gives the reverse inclusion. □

8.2 Products of two commutators

For matrices A, X, Y and scalars a, b, the shared-factor expression is

$$a[A, X][A, Y] + b[A, Y][A, X].$$

Proposition B.4 gives its image for $n \geq 3$.

Proposition 8.5. *Suppose p is in the six-product span of Lemma 8.1 and $n \geq 3$. If it is traceful, its image is $M_n(K)$. If it is trace-zero, its image is the traceless subspace unless it is a scalar multiple of s_4.*

Proof. Write $\alpha, \beta, \gamma, \delta, \epsilon, \zeta$ for the coefficients in the displayed order of the six-product span. In the trace calculation the two opposite orders of a product contribute the same trace. Thus the trace criterion is $(\alpha + \beta, \gamma + \delta, \epsilon + \zeta) \in K(1, -1, 1)$. The restrictions to the three

input tuples $(A,B,A,C),(A,B,C,A),(A,A,B,C)$ give the shared-factor expression with coefficient pairs

$$(\alpha-\zeta,\ \beta-\epsilon),\qquad(-\alpha-\delta,\ -\beta-\gamma),\qquad(\gamma+\zeta,\ \delta+\epsilon),$$

respectively. If p is traceful, the sums of these pairs cannot all vanish, by the trace criterion. The shared-factor theorem gives surjectivity. If the trace is zero, each pair has sum zero. If one pair is nonzero, the same theorem gives all traceless targets. If all pairs vanish, the coefficients are $\alpha(1,1,-1,-1,1,1)$. Expansion gives αs_4. This proves the proposition. □

Corollary 8.6. *Over any field, a proper quartic with nonzero formal trace is surjective in every order at least three.*

Proof. If some $q_i \neq 0$, apply Proposition 8.3. If all vanish, the polynomial is in the six-product span by Lemma 8.1, so Proposition 8.5 applies. □

9 Two-state evaluations of quartics

Let p be a multilinear quartic such that each substitution of the identity for one variable has zero formal trace. Fix two distinct slots X,Y, and order the two remaining slots by their original labels. For pairs U,V,W of scalars, define $H(U,V,W)$ by summing the word coefficients for which X precedes Y, multiplying by the coordinate of U, V, or W for each auxiliary letter according as it occurs before X, between X,Y, or after Y. Define $G(U,V,W)$ with X,Y interchanged. For example, if the remaining slots are Z_1,Z_2, the word Z_2XYZ_1 contributes its coefficient times U_2W_1 to H.

For auxiliary states $u,v\in K^2$, let κ be the sum of the cyclic coefficients for the classes in which Y immediately follows X in cyclic order. Put

$$\begin{aligned} r&=H(u,v,u),\quad s=G(u,v,u),\quad q=G(u,u,v),\quad h=H(v,u,u),\\ q'&=G(v,v,u),\quad h'=H(u,v,v),\qquad t=\kappa(u_1-v_1)(u_2-v_2).\end{aligned}$$

The two-state trace relations are

$$H(v,u,v)=t-s,\qquad G(v,u,v)=t-r. \tag{12}$$

Indeed, $H(U,V,U)+G(V,U,V)$ groups its terms by the cyclic word classes, assigning the state V to the auxiliary letters on the cyclic arc from X to Y and U to those on the other arc. Identifying $U_i=V_i$ deletes the distinction for that letter and gives zero by the corresponding unit substitution. The expression therefore is a scalar times $(U_1-V_1)(U_2-V_2)$, with coefficient of U_1U_2 equal to κ. To see that at least one κ is nonzero when p has nonzero formal trace, let $(c_1,\dots,c_6)$ be the cyclic coefficients in the order

$$1234,\quad 1243,\quad 1324,\quad 1342,\quad 1423,\quad 1432.$$

The coefficients of one cyclic class after deleting $4,3,2,1$, respectively, give

$$c_1+c_2+c_5=c_1+c_2+c_3=c_1+c_3+c_4=c_1+c_4+c_5=0.$$

The sum of both classes after any deletion gives $\sum_i c_i=0$. Thus $c_3=c_5=-c_1-c_2$, $c_4=c_2$, and $c_6=c_1$. For the pairs $(1,2),(1,3),(1,4)$, the resulting sums are $c_1+c_2,-c_1,-c_2$.

Proposition 9.1. *Assume every scalar matrix of size three is a value of p. At some auxiliary states u,v, suppose $t\neq 0$ and one of the following conditions holds:*

1. *$qh \neq 0$ and $(r, s) \neq (0, 0)$.*
2. *$qq'r \neq 0$ or $hh's \neq 0$.*
3. *The coefficient sum of p is zero, $q \neq 0$, $h = r = 0$, $s = t$, and $(H(v, v, u), G(v, u, u)) \neq (0, 0)$, or the corresponding condition for the reversed polynomial.*

Then p is surjective in all sizes at least two.

Proof. Freeze the auxiliary slots at the block diagonal matrices $\mathrm{diag}(u_i I, v_i I)$, with block dimensions specified below. When X has zero diagonal blocks, put $X_{12} = U$, $X_{21} = V$, $Y_{12} = x$, $Y_{21} = y$. Write p_{ij} for the (i, j) block of the output. Expanding the words by their three auxiliary states gives

$$p_{11} = rUy + sxV, \qquad p_{22} = (t - s)Vx + (t - r)yU.$$

For the first condition set $Y_{22} = 0$, $Y_{11} = q^{-1}D$. The off-diagonal blocks are DU, ηVD, where $\eta = h/q$. For equal block sizes use $U = D = I$, $x = 0$, $y = A$, $V = \eta^{-1}B$. This gives

$$\begin{pmatrix} rA & I \\ B & (t - r)A \end{pmatrix}.$$

It gives a nonscalar value and a value of nonzero trace in size two, so Theorem 6.2 supplies all size-two targets. Lemma 7.2 gives every even size. For block sizes $j + 1, j$, use J_j, C_j from Lemma 7.3. Let S be the cyclic permutation matrix on the first block sending the i-th coordinate vector to the $(i + 1)$-st for $1 \leq i \leq j$ and the last to the first. Then $SJ_j = C_j$. If $r \neq 0$, choose $U = J_j$, $y = [A, w]$, $x = 0$, $D = S$, $V = \eta^{-1}BS^{-1}$, where w is a column of length j. This gives

$$\begin{pmatrix} rJ_j[A & w] & C_j \\ B & & (t - r)A \end{pmatrix}.$$

If $r = 0$, use $V = J_j^{\mathsf{T}}$, $x = [A \;\; w]^{\mathsf{T}}$, $y = 0$, $D = \eta^{-1}S^{\mathsf{T}}$ and $U = D^{-1}B^{\mathsf{T}}$. This gives the transpose of the same form with s in place of r. Lemma 7.3 supplies the odd cyclic classes and a scalar singleton plus any even cyclic block of size at least four.

For the second condition it suffices to handle $qq'r \neq 0$, since evaluating the reversed polynomial at the transposes gives the transpose of the original value. Reversal exchanges r, s, q, h, and q', h'. If $h \neq 0$, the first condition applies. If $h = 0$, take $X_{21} = (q')^{-1}V$, $Y_{12} = q'x$ in place of the previous blocks and use $Y_{22} = I$, $Y_{11} = q^{-1}(D - h'I)$. The diagonal products $Y_{12}X_{21}$ and $X_{21}Y_{12}$ are unchanged by these two scalings. In the 12 block the terms are $qY_{11}U + h'UY_{22} = DU$. In the 21 block they are $hX_{21}Y_{11} + q'Y_{22}X_{21} = V$. These give the same equal-size and odd-size sections, using $V = B$ and the preceding $r \neq 0$ choices for the other blocks.

For the third condition retain the four off-diagonal blocks U, V, x, y. Set $X_{11} = cI$, $X_{22} = dI$, $Y_{22} = 0$, $Y_{11} = D_0$, where $c, d \in K$ and D_0 has the first block's order. The coefficients of x and y from the diagonal blocks of X are

$$\alpha = cH(u, u, v) + dG(u, v, v), \qquad \beta = dH(v, v, u) + cG(v, u, u).$$

Choose c, d with $\beta \neq 0$. Terms entirely within one diagonal block vanish by the coefficient sum assumption. The terms from X_{12}, X_{21} give $txV, tyU, qD_0U, 0$ in the blocks $11, 22, 12, 21$ by $r = h = 0$, $s = t$, and (12). The terms from X_{11}, X_{22} in the off-diagonal blocks give $\alpha x, \beta y$. Thus the value is

$$\begin{pmatrix} txV & \alpha x + qD_0U \\ \beta y & tyU \end{pmatrix}.$$

For equal blocks put $U = \beta I$, $y = \beta^{-1} I$, $x = A$, $V = I$ and $D_0 = (q\beta)^{-1}(B - \alpha A)$. The output is $\left(\begin{smallmatrix} tA & B \\ I & tI \end{smallmatrix}\right)$. In a basis of a half-dimensional space and its images, the target has form $\left(\begin{smallmatrix} 0 & B_0 \\ I & C_0 \end{smallmatrix}\right)$. For $S = \left(\begin{smallmatrix} I & C_0 - tI \\ 0 & I \end{smallmatrix}\right)$ and this target matrix T_0,

$$ST_0S^{-1} = \begin{pmatrix} C_0 - tI & B_0 + t(C_0 - tI) \\ I & tI \end{pmatrix}.$$

This has the required form with $A = (C_0 - tI)/t$. By Lemma 7.1, the only remaining even-size pieces are scalar pairs. They come from the size-two theorem, since the displayed evaluations include nonscalar and nonzero-trace values. For dimensions $j + 1, j$, take $V = J_j^{\mathsf{T}}$, $x = [A \quad w]^{\mathsf{T}}$, $y = \beta^{-1} C_j^{\mathsf{T}}$, and $U = \beta C_j$. Choose the last j columns of D_0 to be $(B^{\mathsf{T}} - \alpha x)/(q\beta)$, so $q\beta D_0 C_j = B^{\mathsf{T}} - \alpha x$. The output is the transpose of $\left(\begin{smallmatrix} tJ_j[A & w] & C_j \\ B & & tI_j \end{smallmatrix}\right)$. The fixed bottom-block version of Lemma 7.3 applies.

It remains to obtain noncyclic nonscalar size-three matrices. Each has form $\lambda I_3 + R$, with R of rank one and trace $\mu - \lambda$, by rational canonical form. For fixed λ, μ all such matrices are similar, since rank-one operators of given trace are similar. If $\lambda\mu = 0$, a size-two block and a zero singleton suffice: the rank-one operator splits off a zero line if $\lambda = 0$, and the zero eigenspace splits off a line if $\lambda \neq 0$ and $\mu = 0$.

Under the first condition, for $r \neq s$ use $U = (1, \lambda)^{\mathsf{T}}$, $V = (0, 1)$, $D = \left(\begin{smallmatrix} 1 & 0 \\ -\mu\lambda/\eta & \mu/\eta \end{smallmatrix}\right)$, $y = AV$, $x = BU$, where $rA + sB = 1$ and $(t - r)A + (t - s)B = (\lambda + \mu)/\lambda$. The determinant of the two equations is $t(r - s)$. The resulting rank-one difference from λI_3 is $(1, 0, \mu)^{\mathsf{T}}(-\lambda, 1, 1)$. For $r = s = a$, put $b = (a/t)(2\lambda + \mu) - \lambda$ and $c = \mu - b$. If $bc \neq 0$, take

$$U = \frac{\eta}{bc}\begin{pmatrix} 1 \\ \lambda \end{pmatrix}, \quad V = (1, 0), \quad x = \begin{pmatrix} b/a \\ 0 \end{pmatrix}, \quad y = (0, bc/(a\eta)), \quad D = \begin{pmatrix} c(b - \lambda)/\eta & c/\eta \\ 0 & 0 \end{pmatrix}.$$

The difference is $(1, 0, c)^{\mathsf{T}}(b - \lambda, 1, 1)$. If $b = 0$ the first construction works with $A = 1/a$, $B = 0$. If $c = 0$, use $U = (1, 0)^{\mathsf{T}}$, $V = (0, 1)$, $x = (0, \lambda/a)^{\mathsf{T}}$, $y = (\mu/a, 1/a)$, $D = \operatorname{diag}(1, 0)$. The output is $\left(\begin{smallmatrix} \mu & 1 & 1 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{smallmatrix}\right)$.

Under the second condition with $h = 0$, take $U = (0, 1)^{\mathsf{T}}$ and $V = (1, 0)$. If $s \neq 0$, put $a = (s/t)(2\lambda + \mu) - \lambda - (s/r)\lambda$, $b = \mu - \lambda - a$, and set

$$x = \begin{pmatrix} (a + \lambda)/s \\ 1/s \end{pmatrix}, \quad y = (0, \lambda/r), \quad D = \begin{pmatrix} 0 & ab \\ 0 & b \end{pmatrix}.$$

The difference is $(a, 1, 1)^{\mathsf{T}}(1, 0, b)$. If $s = 0$, put $a = (\lambda + \mu - (t - r)\lambda/r)/t$ and use $x = (a, 0)^{\mathsf{T}}$, $y = (1/r, \lambda/r)$, $D = \left(\begin{smallmatrix} 0 & -\lambda\mu \\ 0 & \mu \end{smallmatrix}\right)$. The difference is $(-\lambda, 1, 1)^{\mathsf{T}}(1, 0, \mu)$.

Under the third condition set $x = (0, 1)^{\mathsf{T}}$, $V = (1/t, \lambda/t)$, $y = (\beta^{-1}, 0)$, $U = (\beta(\lambda + \mu)/t, 1)^{\mathsf{T}}$ and $D_0 = \left(\begin{smallmatrix} 0 & -\lambda\mu/q \\ 0 & (\mu - \alpha)/q \end{smallmatrix}\right)$. The same rank-one difference $(-\lambda, 1, 1)^{\mathsf{T}}(1, 0, \mu)$ results. All these identities follow by substitution into the displayed block formulas.

Now size three is complete. In odd size, decompose the target into cyclic blocks over K. If one odd block has size at least three, use the odd cyclic section for it. Otherwise combine a singleton with an even block of size at least four, if one exists. In the remaining case use one singleton and a two-dimensional block, or three singletons. Each complement has even size, so direct sums of the original input tuples supply the whole target. □

10 Scalar values and auxiliary states

Lemma 10.1. *A multilinear quartic with nonzero formal trace and with zero formal trace in all single-identity substitutions attains I_3 over its coefficient field.*

Proof. Choose a pair with $\kappa \neq 0$ as in Section 9 and scale p so that $\kappa = 1$. Write $r(U,V) = H(U,V,U)$ and $\Delta(U,V) = (U_1 - V_1)(U_2 - V_2)$. Let $E = K^3$ with σ cyclically permuting the factors if K is infinite. For $K = \mathbf{F}_q$ finite, take $E = \mathbf{F}_{q^3}$ and $\sigma : z \mapsto z^q$. Write M_a for multiplication by a on the three-dimensional K-space E. In the slots X, Y, Z_1, Z_2 put

$$\sigma, \qquad M_y\sigma^{-1}, \qquad M_z, \qquad M_w,$$

respectively. Put $R = r((z,w),(\sigma z,\sigma w))$ and $D = (z - \sigma z)(w - \sigma w)$. Using $\sigma M_a = M_{\sigma(a)}\sigma$ and (12), the output is $M_{L(y)}$, where

$$L(y) = R\sigma(y) + \sigma^{-1}(D - R)y.$$

Indeed the words with X before Y contribute $R\sigma(y)$, while those with Y before X contribute $G((z,w),(\sigma^{-1}z,\sigma^{-1}w),(z,w))y$.

Over a field that splits E into three factors, the determinant of L is

$$\prod_{(P,Q)\in\{(U,V),(V,W),(W,U)\}} r(P,Q) + \prod_{(P,Q)\in\{(U,V),(V,W),(W,U)\}} \big(\Delta(P,Q) - r(P,Q)\big), \tag{13}$$

where U, V, W in the cycle are the three pairs of coordinates of z, w. This follows by using the diagonal entries or the cyclic permutation entries of L. Each factor has degree one in the coordinates of z and one in those of w. The invertible change from the split coordinates to the coordinates of a K-basis of E preserves these separate degrees, so the determinant in the K-coordinates has individual degrees at most three. If this polynomial is nonzero and $|K| \geq 4$, it has a nonzero value on K^6. Indeed a polynomial of degree at most three in one variable with four distinct roots is zero, and successively applying this fact to the coefficient polynomials proves the assertion in all six variables. Solving $L(y) = 1$ then gives I_3.

Suppose the determinant polynomial vanishes. Let A be the sum of the four coefficients of r. At the cyclic states U, U, V the determinant is

$$AU_1U_2\Delta(U,V)\big(r(U,V) + r(V,U) - \Delta(U,V)\big).$$

If $A \neq 0$, the parenthesis must vanish. The coefficients of U_1U_2 and V_1V_2 in it are the sum of the within-state coefficients of r minus 1. The coefficients of U_1V_2 and V_1U_2 are the sum of the mixed coefficients plus 1. Both sums therefore add to zero, contrary to $A \neq 0$. Write

$$r = \alpha\Delta + \beta(U_1 - V_1)V_2 + \gamma V_1(U_2 - V_2)$$

using $A = 0$. Define the four expressions

$$1 - 3\alpha + 3\alpha^2 - \beta^2 - \gamma^2, \qquad \beta(2\alpha - 1 - \beta), \qquad \gamma(2\alpha - 1 - \gamma), \qquad \beta\gamma.$$

The coefficients of the determinant for the exponent vectors

$$(2,2,1,1,0,0),\ (2,2,1,0,0,1),\ (2,2,0,1,1,0),\ (2,1,1,1,0,1)$$

in $(U_1, U_2, V_1, V_2, W_1, W_2)$ are the combinations of these expressions with rows

$$(1,-1,-1,0), \quad (-1,2,1,-1), \quad (-1,1,2,-1), \quad (0,-3,0,2).$$

These rows follow by expanding the two triple products in (13). To obtain these rows, write a, b, c, d for the coefficients of $U_1U_2, U_1V_2, V_1U_2, V_1V_2$ in r. The monomial products at the four exponents in the first triple product contribute $a^2d, abd, acd, a^2c + b^2d$, respectively. In the second triple product replace (a,b,c,d) by $(1-a,-1-b,-1-c,1-d)$. Substituting

$(a,b,c,d) = (\alpha, \beta-\alpha, \gamma-\alpha, \alpha-\beta-\gamma)$ and adding the contributions gives the displayed rows. Their determinant is -1, so the four expressions vanish. If $\beta = \gamma = 0$, then $r = \alpha\Delta$. Otherwise one of β, γ is nonzero and the other is zero. For example, if $\beta \neq 0$, then $\beta = 2\alpha - 1$ and $\alpha(1-\alpha) = 0$. Thus, up to exchanging the two coordinates, the two remaining possibilities are $r = U_1(U_2 - V_2)$ and $r = -V_1(U_2 - V_2)$.

Choose w with $\delta = w - \sigma(w)$ a unit, by taking distinct coordinates in the split case or $w \notin K$ in the finite-field case. When $r = \alpha\Delta$, also choose z with $z - \sigma(z)$ a unit and set $y = \sigma^{-1}(D)^{-1}$. Then $L(y) = 1$. For $r = U_1(U_2 - V_2)$, putting $v = \sigma^{-1}(\delta)y$ gives $L(y) = z(\sigma(v) - v)$. Choose a unit z whose inverse has zero trace in E/K, and solve $\sigma(v) - v = z^{-1}$. For $r = -V_1(U_2 - V_2)$ put $v = z\sigma^{-1}(z)\sigma^{-1}(\delta)y$. Then $L(y) = (v - \sigma(v))/z$, and we choose a trace-zero unit z and solve $v - \sigma(v) = z$. The trace is the sum of the three cyclic transforms. It is nonzero because it is the coordinate sum after scalar extension to the split algebra. The kernel of $\sigma - 1$ consists of the scalars, so its image and the trace kernel are both two-dimensional. The image is contained in the trace kernel by cancellation of the cyclic transforms, hence equals it. A nonzero trace-zero element of the finite extension is a unit. In the infinite split case choose coordinates $1, b, -1-b$ with $b \notin \{0, -1\}$.

If $K = \mathbf{F}_3$ and p is not proper, a nonzero single-identity substitution has image $\mathfrak{sl}_3(K)$ by Theorem 1.3 and hence attains I_3. The proper case follows from Corollary 8.6. For $K = \mathbf{F}_2$, use the following table. Encode $a_0 + a_1\theta + a_2\theta^2$ in $\mathbf{F}_8 = \mathbf{F}_2[\theta]/(\theta^3 + \theta + 1)$ by $a_0 + 2a_1 + 4a_2$. The row label gives the coefficients of r in the order $U_1U_2, U_1V_2, V_1U_2, V_1V_2$. Each listed triple (z, w, y) satisfies $Ry^2 + (D - R)^4 y = 1$, using $\theta^3 = \theta + 1$. For the entry 0000, $z = w = \theta$ and $y = 1 + \theta$, so $R = 0$, $D = (\theta + \theta^2)^2 = \theta$, and $D^4 y = (\theta + \theta^2)(1 + \theta) = 1$. For every entry the same reduction replaces each monomial of degree at least three using $\theta^j = \theta^{j-2} + \theta^{j-3}$.

0000	$(2,2,3)$	0001	$(2,2,2)$	0010	$(2,4,5)$	0011	$(2,2,4)$
0100	$(2,3,2)$	0101	$(2,2,4)$	0110	$(2,2,3)$	0111	$(2,2,2)$
1000	$(2,3,2)$	1001	$(2,2,3)$	1010	$(2,3,2)$	1011	$(2,2,2)$
1100	$(3,2,2)$	1101	$(2,2,2)$	1110	$(2,3,3)$	1111	$(2,2,3)$

Rescaling one input removes the normalization of p. □

Lemma 10.2. *Let K have at least three elements. If p has zero formal trace in each single-identity substitution, has nonzero formal trace, and is not proper, then it has a pair and auxiliary states satisfying one of the first two conditions of Proposition 9.1, with $t \neq 0$.*

Proof. The coefficient calculation in Appendix A gives a pair with $\kappa \neq 0$ and both q, h nonzero as polynomials, after exchanging the two slots if needed. Normalize $\kappa = 1$, $u_i = z_i + 1$, $v_i = z_i$. The substitution identifies the space of polynomials which use exactly one of U_i, V_i for each i with the space of polynomials of individual degrees at most one in z_1, z_2. This follows from the independent polynomials $z_i + 1, z_i$ in each variable. Thus q, h remain nonzero, and at least one of r, s is nonzero by (12). For $|K| \geq 4$, the product of q, h and a nonzero one of r, s has individual degrees at most three and is nonzero at some K-point. The first condition holds there.

For $\mathbf{F}_3$ put $\iota(z_1, z_2) = (-z_1 - 1, -z_2 - 1)$. Exchanging u, v after normalization multiplies each auxiliary coordinate by -1, so $q' = q \circ \iota$, $h' = h \circ \iota$ and $s = 1 - r \circ \iota$. In the coordinates $z = (x + 1, y + 1)$, the map ι is $(x, y) \mapsto (-x, -y)$. Let Q and H_0 be the sets of points where q and h are nonzero, and set $S = Q \cap H_0$. The product qh has individual degrees at most two, so S is nonempty. If neither condition holds anywhere, then $r(z) = s(z) = 0$ for $z \in S$, so $r(-z) = s(-z) = 1$. The second condition forces $q(-z) = h(-z) = 0$. Counting $-S$ in the common zero set gives $|Q| + |H_0| \leq 9$.

A nonzero polynomial of individual degrees at most one on $\mathbf{F}_3^2$ has support size $4, 6, 7$, or 9. In fact, in $(dx + c)y + (bx + a)$, if $d \neq 0$, two rows have two nonzeros each, and the

exceptional row has zero or three. If $d = 0$ the nonconstant affine cases have six nonzeros and the constant case nine. Thus both supports have size four, and up to nonzero factors $q = (x - a)(y - b)$ and $h = (x - c)(y - d)$. If $a = c$, one of the two permitted x values has negative distinct from a. If $b \neq d$, there is no common zero on that negative row, and if $b = d$ the two permitted y values cannot both negate to b. Both contradict the property of $-S$. Thus $a \neq c$, and interchanging coordinates gives $b \neq d$. The single point of S is $(-a - c, -b - d)$, whose negative lies at (a, d) or (c, b). In the first case this forces $c = b = 0$ and $a, d \neq 0$. In the second case it forces $a = d = 0$, and exchanging coordinates reduces to the first case. On $(0, \pm 1)$ both $q(z), q(-z)$ are nonzero, so failure of the second condition forces $r(0, \pm 1) = 0$. On $(\pm 1, 0)$ both $h(z), h(-z)$ are nonzero, so $r(\pm 1, 0) = 1$. But

$$r(0, 1) + r(0, -1) = 2r(0, 0) = r(1, 0) + r(-1, 0),$$

a contradiction in $\mathbf{F}_3$. □

11 The remaining quartics over $\mathbf{F}_2$

We continue to assume that the quartic has nonzero formal trace, that each identity substitution has zero formal trace, and that the quartic is not proper. The coefficient sum is zero. A mask means the integer whose bit i is the coefficient of word i in lexicographic order on the 24 permutations of $(1, 2, 3, 4)$, indexed from zero. A mask and its bitwise complement differ by s_4, the signed sum of all 24 words. In characteristic two all signs in s_4 equal one.

Proposition 11.1. *Under these hypotheses every quartic over* $\mathbf{F}_2$ *is surjective in orders* $n \geq 3$.

Proof. Appendix A gives the coefficient cover. Apart from the conditions of Proposition 9.1, the cover consists of 24 representatives and their complements, up to permuting the slots and reversing the words. For six representatives a repeated-slot substitution gives $X[A^2, B]$ or $[A^2, B]X$. Such substitutions kill s_4. Represent $E = \mathbf{F}_{2^n}$ on its underlying $\mathbf{F}_2$-space by multiplication operators M_a, and use the Frobenius σ of E. For $a \notin \mathbf{F}_2$ the commutator $[M_a^2, \sigma] = M_{a^2 - a^4}\sigma$ is invertible. Varying X gives surjectivity.

Mask 783 specializes to $A^2[B, C]$ upon identifying the first two variables, as does its complement. Given a target T, choose a nonzero $b \in E$ with $\operatorname{tr}(M_b T) = 0$ by linear algebra, since $\dim_{\mathbf{F}_2} E = n \geq 2$. Choose a with $a^2 = b^{-1}$. By Lemma 2.1, write $M_b T = [B, C]$. The evaluation at $A = M_a$ gives T.

Mask 2600 specializes to $A[A, C]B + B[A, C]A$ by equating the first and third slots, up to renaming the remaining slots, as does its complement. Lemma 11.2 below proves surjectivity. The other sixteen representatives and their complements have the linear evaluations described in Appendix A. We explain their use.

Fix three inputs, keeping the same free slot in every block. For each pair of block sizes, the linear map on the free rectangular block is the sum of left and right products of the corresponding fixed blocks, in their word order. Replacing all fixed size-two inputs by their tensor products with I_h tensors the size-two linear operator with the identity on M_h. Between this region and a fixed block of size one or three, the rectangular maps are direct sums of h copies of the small rectangular maps, by grouping coordinates with the same tensor index.

For the 21 masks with a size-two source and a scalar border, the square map on M_2 and both rectangular maps are invertible. Hence in even size the linear map is invertible, and in odd size $2h + 1$ it attains exactly the matrices whose last diagonal entry is zero. The scalar-to-scalar operator is zero since the coefficient sum vanishes. Every nonscalar target T is similar to such a matrix: choose independent v, Tv and take a complement to Kv containing

Tv. These sources give I_2, and Lemma 10.1 gives I_3. Direct sums and scalar multiplication give every scalar in odd size as well.

For the remaining 11 masks the even-size certificates have the following form. An invertible size-two map is used by tensoring. A rank-three size-two map whose trace annihilator has rank one and trace one has image similar to the hyperplane with lower-right entry zero: the annihilator is an idempotent similar to E_{22}. Tensoring attains the matrices $\left(\begin{smallmatrix} H & I \\ K & 0 \end{smallmatrix}\right)$. Mask 832207 identifies its first two inputs, giving $ABAC + BAAC + CAAB + CABA$. At $A = \operatorname{diag}(I, 0)$, $B = \left(\begin{smallmatrix} 0 & I \\ K & 0 \end{smallmatrix}\right)$ and $C = \left(\begin{smallmatrix} I & 0 \\ H & 0 \end{smallmatrix}\right)$ this is $\left(\begin{smallmatrix} 0 & I \\ K & H \end{smallmatrix}\right)$. Conjugating the transpose by the block exchange gives $\left(\begin{smallmatrix} H^{\mathsf{T}} & I \\ K^{\mathsf{T}} & 0 \end{smallmatrix}\right)$. A matrix is similar to its transpose, so this evaluation also supplies the half-space pieces. The repeated-input section is unchanged by s_4. The size-two operators also agree for complementary masks, since s_4 vanishes on M_2. By alternation and multilinearity, evaluate at $(E_{11}, E_{12}, E_{21}, E_{22})$. The only nonzero products occur in orders $1243, 2431, 3124, 4312$. With signs they are $-E_{11}, E_{11}, E_{22}, -E_{22}$, respectively. Each displayed size-two source has a nonscalar value and a nonzero-trace value. Thus scalar pairs are available by Theorem 6.2, and Lemma 7.1 gives all even sizes.

For these 11 masks there are two further certificates. Each consists of size-two and size-three fixed inputs with invertible rectangular maps between them. In both certificates the size-two image contains $\operatorname{span}\{E_{11}, E_{12}, E_{22}\}$. The size-three image contains, respectively,

$$\operatorname{span}\{E_{11}, E_{12}, E_{32}, E_{21} + E_{13}\}, \qquad \operatorname{span}\{E_{11}, E_{22}, E_{23}, E_{32}\}.$$

In size $2h + 3$ tensor the size-two inputs with I_h and adjoin the size-three inputs. Thus in the $2h$ region every block upper-triangular matrix with equal block sizes h is attained. All entries between this region and the size-three region are free. In size three only the specified size-three image is required.

For a monic polynomial $t^m + a_1 t^{m-1} + \cdots + a_m$, form a matrix with entries

$$\begin{gathered} F_{11} = -a_1, \quad F_{21} = 1, \quad F_{2j-1,2j} = -a_{2j}, \quad F_{2j+1,2j} = -a_{2j+1}, \\ F_{2j-1,2j+1} = 1, \quad F_{2j+2,2j} = 1 \end{gathered} \tag{14}$$

for $j \geq 1$ and indices in range, with other entries zero. In decreasing odd coordinate order followed by increasing even coordinate order, the constant entries form a directed path with all weights one. The coefficient entry $-a_i$ closes the path segment of i vertices from its row coordinate to its column coordinate. All these cycles contain vertex 1, with $-a_1$ giving a loop there. A nonzero permutation term can therefore use at most one such cycle. Its sign $(-1)^{i-1}$ and the i signs from the entries of $tI - F$ give the contribution $a_i t^{m-i}$. The empty choice gives t^m. In the path order the matrix is upper Hessenberg with subdiagonal entries one, so it is cyclic.

For odd $m = 2h + 3$ put coordinates $1, 2, 3$ in the size-three region and order the other coordinates by parity, odd first. The $2h$ region is block upper triangular, since the only edge from an odd coordinate to an even coordinate, apart from same-parity edges, is $1 \to 2$. The size-three region is $-a_1 E_{11} - a_2 E_{12} - a_3 E_{32} + E_{21} + E_{13}$. The first certificate thus attains every odd cyclic class. For a singleton λ plus an even cyclic matrix of size $2h + 2$, put the singleton and the first two coordinates of F in the size-three region. Now the core is $\lambda E_{11} - a_1 E_{22} - a_2 E_{23} + E_{32}$, and the $2h$ region is again block upper triangular. The second certificate supplies this target.

In particular every noncyclic nonscalar size-three class is covered, since it has a singleton and a size-two cyclic summand. Lemma 10.1 supplies the size-three scalars. In larger odd size use a cyclic odd summand of size at least three if available. Otherwise combine a singleton with an even cyclic summand or take three singletons. The complement has even size in each case. Direct sums of the original input tuples and simultaneous similarity then prove the assertion. □

Lemma 11.2. *The map* $(A, B, C) \mapsto A[A, C]B + B[A, C]A$ *is surjective on* $M_n(\mathbf{F}_2)$ *for* $n \geq 2$.

Proof. Denote the expression by $p(A, B, C)$. For $n = 2h$ use $A = \operatorname{diag}(I_h, 0)$, $C = \left(\begin{smallmatrix} 0 & I \\ K & 0 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & 0 \\ H & I \end{smallmatrix}\right)$ to obtain $\left(\begin{smallmatrix} H & I \\ K & 0 \end{smallmatrix}\right)$. The inputs $A = E_{12}$ and $B = C = E_{21}$ give I_2. Lemma 7.1 now gives all even sizes.

In odd order set $N = \sum_{i=1}^{n-1} E_{i+1,i}$, $P = N + E_{1n}$, and $C = \operatorname{diag}(0, 1, 0, 1, \ldots, 0)P^{-1}$. Conjugating the diagonal factor of C by P shifts its entries one position cyclically. Their differences give $[P, C] = I - E_{11}$ in characteristic two. Since $(I - E_{11})P = N$, substitution yields

$$p(P, PV, C) = P[N, V].$$

Put $V_0 = \sum_{i \text{ odd},\ i<n} E_{i,i+1} + \sum_{i \text{ odd},\ i\leq n-2} E_{i,i+2} + E_{n-1,1}$. The commutator with the first sum is $I - E_{nn}$. The commutator with the sum of the other terms is

$$\sum_{\substack{i \text{ odd} \\ i \leq n-2}} (E_{i+1,i+2} - E_{i,i+1}) + E_{n1} = P^{-1}$$

in characteristic two. Therefore $p(P, PV_0, C) = I + N$, and omitting the first sum gives I. Varying the second input by $e_1 w^{\mathsf{T}}$ changes the output by $e_1 w^{\mathsf{T}} N$, since $P[P, C]e_1 = 0$. This attains $I + N + e_1 z^{\mathsf{T}}$ for every row z^{T} with $z_n = 0$. Its characteristic polynomial is

$$(t + 1)^n + \sum_{i=1}^{n-1} z_i (t + 1)^{n-i}.$$

For the determinant, write $s = t + 1$. In $sI_n - N - e_1 z^{\mathsf{T}}$ the first-row cofactor of column i is s^{n-i}, since the first $i - 1$ columns of the corresponding minor give the subdiagonal product $(-1)^{i-1}$ and the remaining diagonal entries give s^{n-i}. The cofactor sign cancels the subdiagonal sign, and minus equals plus in K. The successive images of e_1 under $N + e_1 z^{\mathsf{T}}$ have leading new coordinates $e_2, \ldots, e_n$ with coefficient one, so the matrices are cyclic. The coefficient $z_n = 0$ gives exactly the monic polynomials that vanish at 1.

Next use $A' = P + I$ with the same C. Write $Q = I - E_{11}$. We have $A'Qe_1 = 0$, and $QA' = N + I - E_{11}$ has right kernel spanned by the all-ones vector, as its first row is zero and the other rows equate consecutive coordinates. With $B = \operatorname{diag}(0, 1, 0, 1, \ldots, 0)$ the value is N: the diagonal entries cancel and consecutive diagonal entries of B sum to one. Varying B by $e_1 w^{\mathsf{T}}$ adds $e_1 w^{\mathsf{T}} QA'$, attaining $N + e_1 z^{\mathsf{T}}$ with $\sum_i z_i = 0$. The same first-row expansion with $s = t$ gives $t^n + \sum_{i=1}^{n} z_i t^{n-i}$, and the same cyclic-vector argument applies. Since $\sum_i z_i = 0$, these are exactly the monic polynomials with value one at 1.

We also attain a singleton 1 plus a Jordan block of size $n - 1$ with eigenvalue 1. Take $T = I + E_{31} + \sum_{i=3}^{n-1} E_{i+1,i}$. The centralizer of N is the span of $I_n, N, \ldots, N^{n-1}$, so $P^{-1}T = [N, V]$ for some V if $\operatorname{tr}(N^j P^{-1} T) = 0$ for $0 \leq j < n$, by the trace-pairing argument of Lemma 2.1. For $j = 0$ the trace equals the sum of the first subdiagonal entries of T and the entry T_{1n}, hence is $n - 3 = 0$ in $\mathbf{F}_2$. For $j \geq 1$, use $N^j P^{-1} = N^{j-1} - E_{j1}$. The trace at $j = 1$ is $\operatorname{tr} T - T_{11} = 0$, and for $j \geq 2$ it is zero by lower triangularity. Hence the evaluation at (P, PV, C) gives T.

For an arbitrary odd-size target decompose into rational cyclic summands. If there is an odd summand of size at least three, take it and an even complement. If the odd summands are singletons, a zero singleton or three 1-singletons can be separated and the complement has even size. The remaining case has exactly one singleton 1 and at least one even cyclic summand. In that even summand split off the $(t+1)$-primary Jordan block of size e, leaving a cyclic summand with polynomial g coprime to $t + 1$. If $e = 0$, the singleton and the

even summand together are cyclic. If $e > 0$ is even, use the preceding singleton-plus-Jordan construction and an even complement. If e is odd, the singleton plus the Jordan block has even size, and g has odd positive degree. Its cyclic summand is covered by the odd cyclic construction, or is a zero singleton if $\deg g = 1$. The original tuples can be taken block diagonal throughout. □

12 Linear sources for the standard quartic

Cyclically rotating a quartic word changes its sign and preserves its trace, so every value of s_4 has trace zero. Substituting I_n for one input gives zero, since the four insertions of I_n in each remaining word have alternating signs. When two inputs coincide, pairing the permutations that exchange those inputs gives zero.

Theorem 12.1. *For every field K and every $n \geq 3$, the image of s_4 on $M_n(K)$ is $\mathfrak{sl}_n(K)$.*

Lemma 12.2. *Every traceless matrix of order three is a value of s_4.*

Proof. Take X, Y with zero diagonal, and write x_{ij}, y_{ij} for their entries. Expansion gives

$$s_4(E_{11}, E_{22}, X, Y) = \begin{pmatrix} 0 & x_{13}y_{32} - x_{32}y_{13} & x_{23}y_{12} - x_{12}y_{23} \\ x_{31}y_{23} - x_{23}y_{31} & 0 & x_{21}y_{13} - x_{13}y_{21} \\ x_{32}y_{21} - x_{21}y_{32} & x_{12}y_{31} - x_{31}y_{12} & 0 \end{pmatrix}.$$

The positions $(1,2),(2,3),(3,1)$ use only the entries of X, Y in $(1,3),(3,2),(2,1)$. They are the three signed 2-by-2 minors of the matrix with rows (x_{13}, x_{32}, x_{21}) and (y_{13}, y_{32}, y_{21}), in cyclic order. The positions $(1,3),(2,1),(3,2)$ use the disjoint triples (x_{23}, x_{12}, x_{31}) and (y_{23}, y_{12}, y_{31}) in the same way. Every vector $w \in K^3$ is the cross product of two vectors, where the cross product has coordinates $u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1$. For $w = 0$, take $u = 0$. For $w \neq 0$, choose $u \neq 0$ with $u \cdot w = 0$, using the coordinate dot product. The map $v \mapsto u \times v$ has kernel Ku by the minor equations and image the plane orthogonal to u by rank-nullity. Thus both triples of output entries can be prescribed independently.

A nonscalar traceless T is similar to a matrix with zero diagonal. Choose v such that v, Tv are independent. Such a vector exists because an operator preserving every line is scalar. Indeed, it is diagonal in any basis, and preservation of the line through a sum of two basis vectors makes their diagonal entries equal. Take a complement to Kv containing Tv. The first diagonal entry is zero and the lower-left column b is nonzero. If the lower-right block is nonscalar, proceed by induction on its order, changing basis only on that block. A traceless block of order one is zero. If it equals λI_r with $\lambda \neq 0$, then $r \geq 2$ since this block has trace zero. Choose a nonzero row f with $fb = 0$. For $S = \left(\begin{smallmatrix} 1 & f \\ 0 & I_r \end{smallmatrix}\right)$, the first diagonal entry of $S^{-1}TS$ is $-fb = 0$, and the lower-right block is $\lambda I_r + bf$. The nonzero rank-one matrix bf makes it nonscalar, so induction applies. A zero lower-right block already has zero diagonal.

The nonzero scalar case occurs only in characteristic three. Expansion gives

$$s_4(E_{12}, E_{23}, E_{31}, \operatorname{diag}(d_1, d_2, d_3)) = \operatorname{diag}(d_2 - d_3, d_3 - d_1, d_1 - d_2).$$

The choice $(d_1, d_2, d_3) = (0, 2\lambda, \lambda)$ gives λI_3. □

12.1 Rectangular maps and sparse companions

For triples $P = (A, B, C)$ in order r and $Q = (A', B', C')$ in order s, let $L_{P,Q}$ send an r-by-s matrix D to the upper right block of

$$s_4\left(A \oplus A',\, B \oplus B',\, C \oplus C',\, \begin{pmatrix} 0 & D \\ 0 & 0 \end{pmatrix}\right).$$

Here $\oplus$ is block diagonal sum. With a block diagonal triple whose blocks are P_i, the output in block (i,j) is $L_{P_i,P_j}(D_{ij})$, since each word uses the fourth argument once. If 0 denotes a zero triple, the maps $L_{P,0}$ and $L_{0,P}$ are left multiplication by $s_3(A,B,C)$ and right multiplication by $-s_3(A,B,C)$, respectively.

Lemma 12.3. *A hollow matrix, meaning a matrix with zero diagonal, is an s_4 value if its coordinates can be partitioned into four classes with zero within-class target blocks, the first three classes having even cardinalities. In particular, every hollow matrix of order at least five supported in positions $|i-j| \le 2$ is such a value.*

Proof. Use the following triples in order two:

$$P_1 = (E_{11}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}), \qquad P_2 = (E_{11}, E_{22}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}), \qquad P_3 = (E_{11}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, I_2).$$

List the entries across each row in order, starting with the first row. In these row-major coordinates, expansion gives the following matrices for $L_{P_1,P_2}, L_{P_1,P_3}, L_{P_2,P_3}$:

$$\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & -2 & -1 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 2 & -2 \\ 0 & -2 & 0 & 1 \\ 0 & 2 & -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & -1 \\ -1 & -2 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Each determinant is one. For rectangular matrices D of size r-by-s and E of size s-by-r, the pairing $(D,E) \mapsto \operatorname{tr}(DE)$ is nondegenerate, since matrix units pair with their transposes. The sum $\operatorname{tr}(s_4(A,B,C,D)E)$ uses one representative with E last from each cyclic class of five distinct labels. Signs are constant on each class because a cyclic rotation has positive sign. Exchanging D,E and rotating each word to put the new last label at the end therefore negates the sum. Apply this identity to the block diagonal triple for P,Q and opposite rectangular inputs D,E. It gives $\operatorname{tr}(L_{P,Q}(D)E) = -\operatorname{tr}(DL_{Q,P}(E))$, so the reverse maps are invertible. The cubic values for P_1, P_2 are $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and that for P_3 is its negative.

Pair the coordinates of the i-th even class, put P_i on each pair, and put the zero triple on the fourth class. Solve each cross-class fourth-input block using the invertible rectangular maps and set the within-class blocks to zero. This constructs the whole target in one evaluation.

For a matrix supported on $|i-j| \le 2$, take the first three classes by residue modulo three. From each odd class remove one coordinate and place it in the fourth class. For the six residues $0, 1, \ldots, 5$ of the order modulo six, choose respectively

$$\varnothing, \quad \{1\}, \quad \{1,5\}, \quad \{1,5,9\}, \quad \{2,6\}, \quad \{3\}.$$

For orders at least five these sets exist, each contains one coordinate from every residue class of odd cardinality and none from the other classes, and any two of its coordinates have distance at least three. □

The Fiedler companion [7] in (14) has bandwidth two (its entries vanish when $|i-j| > 2$), and only its first diagonal entry can be nonzero. Lemma 12.3 therefore gives all cyclic traceless targets of order at least five.

Lemma 12.4. *If C is cyclic of order at least two and $\operatorname{tr} C = -a$, the matrix $a \oplus C$ is similar to a hollow matrix supported on $|i-j| \le 2$.*

Proof. Use the companion above for C. In the order with the scalar first, the leading three-by-three block is $R = \operatorname{diag}(a, -a, 0) + E_{32} + bE_{23}$ for a coefficient b. The matrix $S = \left(\begin{smallmatrix} 1 & 0 & a \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$ gives

$$S^{-1}RS = \begin{pmatrix} 0 & -a & a^2 \\ -a & 0 & b - a^2 \\ 1 & 1 & 0 \end{pmatrix}.$$

The connections to later coordinates are supported in E_{24}, E_{43}, E_{53}, with terms beyond the matrix order omitted. These are unchanged because $S^{-1}e_2 = e_2$ and $e_3^{\mathsf{T}}S = e_3^{\mathsf{T}}$. □

Lemma 12.5. *A traceless direct sum of at most three cyclic matrices of total order at least eight is an s_4 value. The same holds for three cyclic matrices each of order at least two.*

Proof. Put $P_0 = (E_{12}, E_{23}, E_{31} + I_3)$ on a three-dimensional coordinate block. Its fourth-slot map on diagonals gives every trace-zero diagonal by the expansion in Lemma 12.2, since the added identity contributes zero. Use the triples

$$Q_1 = (E_{11}, E_{12} + E_{21}, I_2), \quad Q_2 = (E_{11}, E_{22}, E_{12} + E_{21}), \quad Q_3 = (E_{12}, E_{21}, I_2 + E_{21})$$

on pairs outside that block. The cubic determinants of P_0, Q_1, Q_2, Q_3 are $1, 1, 1, -1$, respectively. For an ordered pair of triples $P = (A_1, A_2, A_3)$, $Q = (B_1, B_2, B_3)$, in row-major order the column of $L_{P,Q}$ for the rectangular unit E_{uv} is obtained by flattening the sum of

$$\operatorname{sgn}(\sigma)\, A_{\sigma(1)} \cdots A_{\sigma(j-1)}\, E_{uv}\, B_{\sigma(j+1)} \cdots B_{\sigma(4)}, \qquad \sigma(j) = 4,$$

over the 24 permutations, with empty products the appropriate identity. The six determinants for pairs in the displayed order are all one, as computed in Appendix C. The reverse maps are invertible by trace duality.

Take the first coordinate from each Fiedler block as part of the three-dimensional block. For three summands this gives a traceless diagonal. For two summands of total order at least eight, choose the fourth coordinate of the larger summand as the third coordinate of the block. It has zero diagonal and no connection to its own first coordinate. The outside coordinates have zero diagonal and retain bandwidth two when the surviving coordinates of each summand are listed in their original order and the resulting lists are concatenated. Here an edge means that at least one of the two entries joining distinct coordinates can be nonzero. If at least five coordinates remain outside, the partition in Lemma 12.3 applies with the triples Q_i replacing P_i. When there are three summands of orders at least two and total order six or seven, the outside consists of three isolated vertices, or one edge and two isolated vertices. In the first case use only the zero class. In the second, pair the two edge endpoints separately with the isolated vertices and give the two pairs different colors. Solve the fourth-input blocks as before. For one summand the cyclic result already applies. □

12.2 Weighted shifts

For arbitrary A, B, C, D, grouping the 24 words of s_4 by the two adjacent pairs gives

$$\begin{aligned} s_4(A, B, C, D) = {} & [A, B][C, D] + [C, D][A, B] - [A, C][B, D] - [B, D][A, C] \\ & + [A, D][B, C] + [B, C][A, D]. \end{aligned}$$

Put $U = [A, C]$, $W = [A, D]$. At $B = A^2$, the first two products vanish and the other four sum to

$$-U(AW + WA) - (AW + WA)U + W(AU + UA) + (AU + UA)W = [A, [U, W]].$$

The same quartic value is obtained with $B = A^2 - 2A + I_n$, by the vanishing identities for s_4. Also $s_3(A, A^2, C) = A[A, C]A$ by grouping its six words, and $s_3(A, I_n, C) = -[A, C]$. Consequently

$$\begin{aligned} s_4(A, A^2 - 2A + I_n, C, D) &= [A, [[A, C], [A, D]]], \\ s_3(A, A^2 - 2A + I_n, C) &= A[A, C]A - [A, C]. \end{aligned} \tag{15}$$

Let $J = \sum_{i=1}^{n-1} E_{i,i+1}$. For $n \geq 3$, take nonzero weights $v_1, \ldots, v_{n-1}$ with sum zero. Put $c_i = \sum_{j=1}^{i} v_j$, with $c_0 = c_{n-1} = 0$, and $C = \sum_{i=1}^{n-2} c_i E_{i+2,i}$. The first lower entries of $[J, C]$ are $c_i - c_{i-1} = v_i$. Thus, with $V = [J, C]$,

$$L(X) = s_4(J, J^2, X, C) = -[J, [V, [J, X]]].$$

A band of offset r is the span of $E_{i,i+r}$ for $1 \leq i, i + r \leq n$. The map L sends offset r to offset $r + 1$, so the outputs from separate input bands can be prescribed independently. Each commutator map is skew-adjoint for $(X, Y) \mapsto \operatorname{tr}(XY)$ by cyclicity of trace. Reversing the three commutator maps preserves their order, since the two outer maps are both $[J, -]$. Hence L is skew-adjoint.

Lemma 12.6. *With these weights, the diagonal and first upper band of L are the trace-zero diagonal and the entire first upper band if $n \neq 0$ in K. If $n = 0$ in K, the additional respective conditions on diagonal entries d_i and first upper entries u_i are*

$$\sum_{i=1}^{n} \ell_i d_i = 0, \qquad \sum_{i=1}^{n-1} i u_i = 0, \qquad \ell_i = \sum_{j=i}^{n-1} j/v_j.$$

The first lower band is given in both cases by the following two conditions on its entries t_i:

$$\sum_{i=1}^{n-1} t_i = \sum_{i=1}^{n-1} w_i t_i = 0, \qquad w_1 = 0, \quad w_{i+1} - w_i = \frac{c_i}{v_i v_{i+1}} \quad (1 \leq i < n - 1).$$

For $3 \leq r \leq n - 1$, the entire upper band of offset r is covered if

$$\sum_{i=1}^{n-r+1} \prod_{j=i}^{i+r-2} v_j \neq 0.$$

Proof. If X is on the first lower band, then $[J, X]$ runs over all trace-zero diagonals, by successive differences. Call its diagonal b_i. The first lower entries of $[[J, X], V]$ are $v_i(b_{i+1} - b_i)$. The outer bracket takes their successive differences to the diagonal, so prescribing d_i determines $b_{i+1} - b_i = (\sum_{j=1}^{i} d_j)/v_i$. The remaining freedom is a common constant. The condition $\sum_i b_i = 0$ is always solvable for $n \neq 0$. If $n = 0$, summing the differences gives

$$\sum_{i=1}^{n} b_i = n b_1 + \sum_{i=1}^{n-1} (n - i)(b_{i+1} - b_i) = -\sum_{j=1}^{n} \ell_j d_j.$$

Here $\ell_n = 0$ by the empty sum convention. For diagonal X, $[J, X]$ runs over all first upper bands. Commuting with V gives all traceless diagonals by taking the differences of the first upper entries multiplied by v_i. Taking the outer bracket then gives precisely the first upper entries $u_i = b_{i+1} - b_i$ for trace-zero b. The additive constant again gives the asserted condition.

For the first lower band, write u'_i for the entries of $[J, X]$ with X in the second lower band. They run over all tuples with sum zero. The second lower entries of $[[J, X], V]$ are

$z_i = u'_{i+1}v_i - v_{i+1}u'_i$. Then $t_i = z_i - z_{i-1}$, with endpoint values $z_0 = z_{n-1} = 0$, so $\sum_i t_i = 0$. The quotients u'_i/v_i have successive differences $z_i/(v_iv_{i+1})$ and can be reconstructed up to a common constant. Multiplying by the weights and summing gives

$$\sum_{i=1}^{n-1} u'_i = \sum_{j=1}^{n-2} \frac{z_j}{v_jv_{j+1}} \sum_{i=j+1}^{n-1} v_i = -\sum_{j=1}^{n-2}(w_{j+1} - w_j)z_j.$$

The common constant contributes zero because $\sum_i v_i = 0$. Since $z_i = \sum_{j=1}^{i} t_j$ and $\sum_i t_i = 0$, the last sum equals $\sum_i w_i t_i$. Thus the two asserted conditions are necessary and sufficient.

For the higher upper bands, it suffices by skew-adjointness to show injectivity on the lower band of offset $-r$. On strictly lower bands $[J, -]$ is injective, again by successive differences with zero boundary entries. If $L(X) = 0$, the entries u'_i of $[J, X]$ satisfy $u'_{i+1}v_i = v_{i+r-1}u'_i$. Thus u'_i is a common multiple of $\prod_{j=i}^{i+r-2} v_j$. Their sum is zero because the entries of $[J, X]$ are successive differences with zero endpoints. Under the stated nonvanishing condition, all are zero and $X = 0$. The trace pairing identifies opposite bands as duals. Thus the rank of L on offset $r - 1$ equals the rank on offset $-r$, namely $n - r$, the dimension of the output band of offset r. □

Lemma 12.7. *Given a monic polynomial of degree $d \geq 3$, prescribe any nonzero first lower entries $t_1, \ldots, t_{d-1}$ in order d. There is a cyclic matrix with that characteristic polynomial, those entries, and zero entries below the first lower band and on the second upper band.*

Proof. Choose diagonal entries d_i with the required sum s and with $d_1 \neq d_3$. Above the diagonal allow entries only at $(1, 2), (2, 3)$ and $(1, j)$ for $4 \leq j \leq d$. The second and third characteristic coefficients, as functions of the entries at $(1, 2), (2, 3)$, have coefficient matrix

$$\begin{pmatrix} -t_1 & -t_2 \\ t_1(s - d_1 - d_2) & t_2(s - d_2 - d_3) \end{pmatrix},$$

whose determinant is $t_1t_2(d_3 - d_1)$. The only off-diagonal terms in the second and third coefficients are the two-cycles on adjacent coordinates, since the second upper band vanishes. An entry in $(1, j)$ with $j \geq 4$ must use the subdiagonal path from 1 to j in any determinant term that contains it. The shortest cycle has length j, is unique and contributes the multiplier $-t_1 \cdots t_{j-1}$ to the j-th coefficient. Hence it sets that coefficient successively. Cyclicity follows because the successive iterates of e_1 have first new coordinates $t_1, t_1t_2, \ldots, t_1 \cdots t_{d-1}$. In order two the same prescription uses only a single upper entry to fix the determinant. □

Lemma 12.8. *Let K be infinite and $n \geq 5$. Every traceless matrix of order n is an s_4 value if $n \neq 0$ in K, or if its minimal polynomial has degree at least four.*

Proof. We first show that the weights in Lemma 12.6 can be chosen with the w_i distinct and every indicated interval sum nonzero. Work on the affine hyperplane $\sum_i v_i = 0$, with $m = n-1$. Each interval sum for offset r is a nonzero polynomial there: specialize $v_1, \ldots, v_{r-1}$ to nonzero entries with sum zero and set the remaining entries to zero. Exactly one product survives. A tuple of a given length at least two with no zero entries and sum zero exists over an infinite field by avoiding the coordinate hyperplanes.

For $1 \leq i < j \leq m$, the rational function $w_j - w_i = \sum_{a=i}^{j-1} c_a/(v_av_{a+1})$ is nonzero. If $i > 1$, its term at $a = i$ has a pole at $v_i = 0$ with nonzero residue c_{i-1}/v_{i+1}, and no other term has that pole. If $i = 1, j < m$, use the pole at $v_j = 0$, with residue c_{j-1}/v_{j-1} from the last term. These residues are nonzero since the prefix and complementary suffix both contain variables. For $i = 1, j = m$, the endpoint terms simplify to $1/v_2$ and $-1/v_{m-1}$. At $v_2 = 0$ the residue of the sum is $1+v_1/v_3$, nonzero because $m \geq 4$. These pole checks hold on

the hyperplane, where distinct coordinate functions are nonproportional. Thus the required simultaneous nonvanishing holds on a nonempty polynomial open set. Since K is infinite, this set has a K-point.

In rational canonical form let h be the number of internal subdiagonal positions in the companion blocks. For $h \geq 3$, the equations $\sum_i t_i = \sum_i w_i t_i = 0$, with t_i supported on those positions, admit a vector nonzero in every internal coordinate. Indeed, after prescribing one coordinate, two other distinct w_i let us solve the equations. Thus each required nonvanishing cuts out the complement of a proper hyperplane in the solution space. When $n \neq 0$, apply Lemma 12.7 to these choices and use Lemma 12.6 on each required band.

For $h \leq 1$, or for $h = 2$ with two quadratic invariant factors, at least one invariant factor has degree one. Divisibility of invariant factors then forces all factors to split in these cases. A Jordan representative has only the diagonal and first upper band, so is attained when $n \neq 0$. The remaining case with $h = 2$ has a cyclic cubic and scalar lines. Put the cubic on coordinates $2, 3, 4$ and give it lower entries $t, -t$ with $t \neq 0$. Choose weights with $v_1 + v_2 = 0$ and a separate nonzero zero-sum tuple on $v_3, \ldots, v_{n-1}$. The latter has length $n - 3 \geq 2$. Then $w_2 = w_3$, which gives the required lower compatibility. The sparse representative uses no second upper band, so it is attained.

Assume now that $n \cdot 1_K = 0$ and put an invariant factor of degree at least four first. The first block can use a third first upper entry at $(3, 4)$. Write $r_i = i/t_i$ for $i = 1, 2, 3$. After factoring out $t_1 t_2 t_3$, the determinant of the three equations for the second and third characteristic coefficients and the first-upper compatibility is, up to sign,

$$g(d) = \det \begin{pmatrix} 1 & 1 & 1 \\ d_1 + d_2 & d_2 + d_3 & d_3 + d_4 \\ r_1 & r_2 & r_3 \end{pmatrix}.$$

Here the first row comes from the second characteristic coefficient, the second from the third after subtracting a multiple of the first, and the third from $\sum_i i u_i = 0$.

In choosing v above also avoid $(3c_2 - v_3)(3v_1 + v_3) = 0$. Each factor is a nonzero polynomial on the hyperplane, including in characteristic three. To show the needed determinant can survive the diagonal conditions, first allow all lower entries except the first three to be zero and take

$$(t_1, t_2, t_3) = (-c_2, c_3, -v_3).$$

This is $v_2 v_3$ times $(w_2 - w_3, w_3 - w_1, w_1 - w_2)$, so it satisfies both lower equations and has its three entries nonzero. The coefficient vector of g on the first four diagonal entries is $(r_2 - r_3, r_2 - r_1, r_3 - r_2, r_1 - r_2)$. Its first-to-second and third-to-fourth differences are $r_1 - r_3$ and $r_3 - r_1$. If it were in the span of the constant vector and $(\ell_i)_{i=1}^4$, then in characteristic three the third difference would have to vanish, whereas it equals $1/c_2$. In other characteristics comparison with $\ell_1 - \ell_2 = 1/v_1$ and $\ell_3 - \ell_4 = 3/v_3$ would force $(r_1 - r_3)(3v_1 + v_3) = 0$, excluded by the choice of weights. A minor witnessing this nonmembership stays nonzero on a polynomial open set after clearing $t_1 t_2 t_3$. Intersect with the nonzero internal coordinate conditions to choose all required t_i.

The affine space of diagonals with the individual block traces and $\sum_i \ell_i d_i = 0$ is nonempty, since $\ell_1 - \ell_2 \neq 0$ permits adjustment inside the first block. The preceding nonmembership implies g is nonzero on that space. For any other block of order at least three, its first and third diagonals may be made unequal by varying them and compensating the ℓ-condition in the first block. Choose all these inequalities and $g \neq 0$ simultaneously. In each other block choose its upper entries as in Lemma 12.7. In the first block the nonsingular three-equation system sets its second and third coefficients and absorbs the global first-upper condition. The entries $(1, j)$ for $j \geq 4$ within each block then set the remaining characteristic coefficients in succession. A determinant cycle containing $(1, j)$ has at least j vertices because the return

from coordinate 1 to j can increase the index only by single steps along the subdiagonal. Thus the added entry $(3,4)$ leaves the coefficient-setting argument of Lemma 12.7 valid. This yields a similar target in the image of L, band by band. □

12.3 Orders four and six

Lemma 12.9. *Every traceless target of order four is an s_4 value. The same holds in order six in characteristic two.*

Proof. First let $2 \neq 0$ in K and work in order four. Put $J = E_{12}+E_{23}+E_{34}$, $C_0 = E_{31}+2E_{42}$, and $C = C_0 + E_{41}$. The map $X \mapsto s_4(J, J^2, X, C_0)$ contains all traceless upper bidiagonal matrices, as well as traceless sums of two blocks of order two whose subdiagonal entries are opposite. Indeed the weights $1, 1, -2$ in Lemma 12.6 have $w_1 = w_3 = 0$. For C in place of C_0 the image is the trace orthogonal complement of I_4, J, J^2, J^3, C, W, where

$$W = \begin{pmatrix} 5 & 0 & 0 & 0 \\ -2 & 3 & 2 & 0 \\ 3/2 & 0 & -1 & 0 \\ -3/2 & 0 & 1 & 0 \end{pmatrix}.$$

For the map $[J, [[J, X], [J, C]]]$, the four powers of J commute with J, the choice $X = C$ gives zero, and direct multiplication gives $[[J, W], [J, C]] = 2J$. Thus all six matrices are in the kernel. They are independent: the $(4,3)$ entry detects W and then the $(4,1)$ entry detects C. The four powers occupy distinct bands. Appendix C computes a rank-ten minor of determinant -64, so the rank is ten whenever $2 \neq 0$. Skew-adjointness now identifies the image with the trace orthogonal complement of this kernel.

Let $p = x^4 + bx^2 + cx + d$ be the characteristic polynomial of a cyclic traceless target. If $p'(a) \neq 0$ for some a, set $t_1 = -p(a)/(2p'(a))$, choose $t_2 \notin \{0, -t_1\}$, and put $t_3 = -t_1 - t_2$. Set

$$z = \frac{2p'(a)}{t_2 t_3}, \qquad u_3 = \frac{3}{2}z - 11a - 2t_2, \qquad u_2 = \frac{-6a^2 - t_3 u_3 - b}{t_2},$$

$$T = \begin{pmatrix} a & 0 & 0 & z \\ t_1 & a & u_2 & -z/2 \\ 0 & t_2 & -3a & u_3 \\ 0 & 0 & t_3 & a \end{pmatrix}.$$

The six trace pairings vanish: the first four follow from the diagonal sum, subdiagonal sum and zeros on lower bands, the pairing with C is $T_{13} + 2T_{24} + T_{14} = 0$, and the pairing with W is $11a + 2t_2 + u_3 - 3z/2 = 0$. Writing $y = x - a$, expansion along the first row gives

$$\det(xI_4 - T) = y^4 + 4ay^3 - (t_2u_2 + t_3u_3)y^2 + \tfrac{1}{2}t_2t_3zy - t_1t_2t_3z.$$

Its x^2-coefficient is b, and its value and derivative at a agree with those of p. These three data determine a monic quartic whose cubic coefficient is zero: the difference is at most linear after fixing the x^2-coefficient and vanishes to order at least two at a. If $t_1 \neq 0$, the successive iterates of the first vector prove cyclicity by the nonzero subdiagonal. Otherwise a is a simple root and the lower block T_0 of order three is cyclic with polynomial coprime to $x - a$. Conjugation $S^{-1}TS$ with $S = \left(\begin{smallmatrix} 1 & R(T_0 - aI_3)^{-1} \\ 0 & I_3 \end{smallmatrix}\right)$ removes the upper-right row R. The direct sum of the two coprime cyclic blocks is cyclic.

If the derivative vanishes at every point, the field has at most three elements because the derivative has degree three with leading coefficient $4 \neq 0$. Thus $K = \mathbf{F}_3$, and its values at $0, 1, -1$ give $c = 0$ and $b = 1$. For $d = 1$, the polynomial is $(x^2 - 1)^2$, so the Jordan form

belongs to the C_0 image. For $d = 0, 2$, respectively, the following matrices X give cyclic outputs with the indicated polynomial under the C map:

$$\begin{pmatrix} 2 & 2 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & 0 & 2 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 0 & 2 \\ 2 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

The outputs, obtained by the commutator formula, are $\left(\begin{smallmatrix} 2 & 2 & 0 & 1 \\ 2 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 2 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 \end{smallmatrix}\right)$. For each output T, the determinant of the columns $e_1, Te_1, T^2e_1, T^3e_1$ is 1 in $\mathbf{F}_3$. Expanding $\det(xI_4 - T)$ gives $x^4 + x^2$ and $x^4 + x^2 + 2$, respectively.

For noncyclic targets in order four, the possible invariant-factor degrees are $(1, 1, 1, 1)$, $(1, 1, 2)$, $(2, 2)$ and $(1, 3)$. The first two cases split by divisibility, and their Jordan forms are covered by C_0. In the $(2, 2)$ case the equal factors are $x^2 + b$ since the total trace is zero and $2 \neq 0$. Use the two zero-diagonal companion blocks with subdiagonal $1, -1$ and upper entries $-b, b$, respectively. These lie together in the C_0 image. In the $(1, 3)$ case the factors are $x - a, (x - a)(x^2 + 2ax + b)$. For $t \neq 0$ use

$$T = \begin{pmatrix} a & u & 0 & 0 \\ t & -3a & u + (b + 3a^2)/t & 0 \\ 0 & -t & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix}, \qquad u = -5a/2 - t.$$

The leading block is cyclic by its nonzero subdiagonal. Expanding its characteristic determinant gives

$$(x - a)\big((x + 3a)(x - a) + b + 3a^2\big) = (x - a)(x^2 + 2ax + b).$$

The first five trace pairings for C vanish by the shape of T, and the W pairing is $-5a - 2u - 2t = 0$.

Now suppose that the characteristic is two. The order-four cyclic case follows from the fixed triple in the first row of Table 4, which contains all hollow order-four Fiedler companions. For a nonscalar order-two matrix the companion with trace τ and determinant δ is $\left(\begin{smallmatrix} \tau & \delta \\ 1 & 0 \end{smallmatrix}\right)$. For matrices $A = (a_{ij})$ and B, write $A \otimes B$ for the block matrix $(a_{ij}B)$. With $H = E_{11}, J = E_{12}, U = E_{21}$ in order two, expansion gives

$$s_4\left(H \otimes I_2,\ (\delta J + U) \otimes I_2,\ H \otimes J,\ \begin{pmatrix} 1 & \tau \\ 0 & 0 \end{pmatrix} \otimes U\right) = \begin{pmatrix} \tau & \delta \\ 1 & 0 \end{pmatrix} \otimes I_2.$$

This covers $(2, 2)$. The fixed triple (E_{12}, E_{23}, E_{34}) in order four gives the respective outputs in the two rows

Input	E_{41}	E_{42}	E_{43}	E_{11}	E_{21}	E_{31}
Output	I_4	E_{12}	E_{13}	E_{14}	E_{24}	E_{34}.

Thus it covers scalar shifts of the nilpotent types $1 + 1 + 1 + 1$, $2 + 1 + 1$, $3 + 1$, using $0, E_{14}, E_{12} + E_{24}$. In degree pattern $(1, 1, 2)$ all eigenvalues are equal, since the quadratic is $(x - a)(x - b)$ and $3a + b = 0$. Pattern $(1, 3)$ has factors $x - a, (x - a)(x^2 + b)$. If $b = a^2$, it gives the type $3 + 1$ after scalar shift. Otherwise it is similar to $aI_2 \oplus C(x^2 + b)$, where $C(q)$ denotes any companion of q. Put (E_{12}, E_{23}, E_{31}) on the first three coordinates and the zero triple on the fourth. Its cubic value on the first block is I_3, since the three nonzero words are E_{11}, E_{22}, E_{33} with positive signs. Hence both rectangular maps are invertible. For this fixed order-three triple, the diagonal input $\mathrm{diag}(0, 0, a)$ gives $\mathrm{diag}(a, a, 0)$ by the formula in

Lemma 12.2. The zero triple gives zero in position $(4,4)$. Prescribing $T_{34} = 1, T_{43} = b$ as the only cross-block entries gives the required similarity type.

For order six in characteristic two, the four cases with a largest invariant factor of degree at least four are the next four rows of Table 4. They contain first-row companions for each block: the entries just below the diagonal are ones, the free first row prescribes the coefficients, and the displayed diagonal terms impose total trace zero. In $(1,1,4)$ the two degree-one factors coincide, and in $(1,5)$ and $(2,4)$ the two block traces agree.

For minimal degree at most three in order six use the next row of the table. It contains one linear space comprising $\mathrm{diag}(R,R)$ for every $R \in M_3(K)$, the leading traceless corner of order two, E_{23}, and the off-diagonal identity blocks $N_+ = \left(\begin{smallmatrix} 0 & I_3 \\ 0 & 0 \end{smallmatrix}\right)$, $N_- = \left(\begin{smallmatrix} 0 & 0 \\ I_3 & 0 \end{smallmatrix}\right)$. Thus it also contains $E_{45} = \mathrm{diag}(E_{12}, E_{12}) + E_{12}$. The possible nilpotent types of largest part at most three have representatives

$$0,\ E_{12},\ E_{12}+E_{45},\ N_+,\ E_{12}+E_{23},\ E_{12}+E_{23}+E_{45},\ \mathrm{diag}(J_3,J_3),$$

where $J_3 = E_{12} + E_{23}$ in order three. The same space contains I_6, so it gives the types whose minimal polynomial is a power of a linear factor. Three irreducible quadratic blocks have trace zero on each block and use $\delta N_+ + N_-$ for determinant δ. Two cubic blocks use $\mathrm{diag}(R,R)$. Split semisimple types have even multiplicities: otherwise two distinct eigenvalues would have sum zero in characteristic two. For minimal polynomial $(x-a)^2(x-b)$ the counts of scalar blocks at a, b are even. With one block of size two at a, start from either bI_6 and change the leading traceless corner, or duplicate $\mathrm{diag}(a,a,b)$ and add E_{12}. With two blocks of size two at a, duplicate $J_2(a) \oplus (b)$, where $J_2(a) = aI_2 + E_{12}$. For a linear factor $x-a$ and irreducible quadratic q, either duplicate $(a) \oplus C(q)$ or change the leading corner of aI_6 by $C(q) - aI_2$; the latter corner has trace zero. These exhaust the factorizations of degree at most three. □

Combining Lemmas 12.2, 12.9 and 12.3 gives all cyclic traceless targets of order at least three. Lemma 12.4 also gives every traceless sum of a scalar line and a cyclic matrix, in total order at least three.

12.4 Scalar shifts of nilpotent matrices

Lemma 12.10. *Every traceless matrix $aI_n + N$, where N is nilpotent and $n \geq 3$, is an s_4 value.*

Proof. If $n \neq 0$ in K, then $a = 0$. Choose weights in Lemma 12.6 by taking all ones in characteristic two, where $n-1$ is even. In other characteristics a zero-sum tuple of nonzero entries of any length at least two can be made from pairs $(1,-1)$ and, when the length is odd, one triple $(1,1,-2)$. The entire first upper band is obtained, including a Jordan representative of N.

In characteristic two with n even, write $n = 4a' + 6b'$ with $b' \in \{0,1\}$. Use a' copies of the order-four triple and, if $b' = 1$, one copy of the order-six triple in the scalar and shift rows of Table 4. Each local fourth-slot image contains the identity and the successive first upper units, as well as E_{41} in the order-four case. The latter supplies the boundary upper unit between identical consecutive blocks of order four. Between the last order-four block and the order-six block, the rectangular map contains E_{41} in the 4-by-6 space, by the last rectangular identity in that group of the table. Thus one fixed triple in order n gives any scalar and first upper band together. This includes the required similarity types.

Suppose $n = mp$ in odd characteristic p. In Lemma 12.6, I_n is covered exactly when $\sum_{i=1}^{n-1} i^2/v_i = 0$. If $m = 1$ or $m \geq 3$, take $v_i = i$ for $p \nmid i$ and a nonzero zero-sum tuple on the $m-1$ remaining entries. For $m = 1$ the latter tuple is empty. Both sums vanish, since

the sum of the nonzero residues in each p-block is zero. For $m = 2, p \geq 5$, take the same $v_i = i$ except for $v_1 = -1, v_p = 8/3, v_{p+1} = 1/3$. Their contribution to the first sum changes by $-2 + 8/3 - 2/3 = 0$, and the nonzero-numerator contribution to the second sum stays $-1 + 3 = 2$. The remaining pair $m = 2, p = 3$ is the order-six characteristic-three scalar and shift row of the table. That row supplies the identity and entire first upper band.

For the weighted choices just made, the first upper entries u_i must satisfy $\sum iu_i = 0$. If N has rank at least two, arrange its Jordan blocks so at least two intended edges are at positions i nonzero in K. A block of length at least three can start at position 1. If all nontrivial blocks have length two and $p \geq 5$, use edges at $1, 3$. If $p = 3$ and there are at least three such blocks, put three first, giving edges at $1, 3, 5$. If $p = 3$ and there are exactly two, the total size is divisible by three, so a singleton can separate them and give edges at $1, 4$. Choose nonzero iu_i on the nonzero-coefficient intended edges summing to zero, using the pairs and triples above. On every other intended edge take $u_i = 1$ and set the unused entries to zero. The resulting weighted Jordan chains are similar to the given ones. The scalar and this band are covered by one fixed triple. Rank zero uses just the diagonal already obtained.

For rank one, use the shift J above and put $H = \sum_{i=1}^{n-1} \binom{i+1}{2} E_{i+1,i}$, with the integers read in K. The diagonal of $[J, H]$ is $(1, \ldots, n)$ because $\binom{i+1}{2} - \binom{i}{2} = i$ and $\binom{n+1}{2} = 0$ in K for odd characteristic dividing n. With the fixed triple (J, J^2, H), the fourth-slot map is $D \mapsto [J, [[J, H], [J, D]]]$, equal to $-(r-1)[J, [J, D]]$ on the input band of offset $r - 2$. Indeed, $[J, D]$ has offset $r - 1$, and commutation with $[J, H]$ multiplies each such entry by $-(r - 1)$. For $r = 2$, the output is the entire second upper band, since one can prescribe the second differences of an input diagonal recursively. For $r = 0$, take $D = \sum_{i=1}^{n-2} \binom{i+1}{2} E_{i+2,i}$. The i-th output diagonal entry is

$$\binom{i+1}{2} - 2\binom{i}{2} + \binom{i-1}{2} = 1.$$

The two coefficients beyond the right end, $\binom{n}{2}, \binom{n+1}{2}$, vanish in K, as do the two left boundary coefficients. By linearity the same slice contains $aI_n + E_{13}$, which gives every rank-one nilpotent shift up to similarity. □

12.5 Targets of small minimal degree

Lemma 12.11. *Suppose K has odd characteristic or characteristic zero, and $T = aI_r \oplus C(q)^{\oplus s}$ is traceless of order $n = r + 2s \geq 5$, where q is an irreducible quadratic and $r, s \geq 1$. Then T is an s_4 value. The same conclusion holds in characteristic two when n is odd.*

Proof. First consider odd characteristic or characteristic zero. If $s = 1$, put the quadratic block on coordinates $2, 3$, with scalar coordinates on $1, 4$. Choose weights with $v_1 + v_2 = 0$ and $v_3 + \cdots + v_{n-1} = 0$, using the nonzero pairs and triples of the preceding proof. Then $w_2 = w_3$. Put the first lower entries at positions $2, 3$ equal to $t, -t$, with $t \neq 0$, so the latter entry couples the block to scalar coordinate 4. Prescribe the quadratic diagonal with sum $\operatorname{tr} C(q)$. If $n = 0$ in K, adjust the difference between the two entries to satisfy the diagonal condition, using $\ell_2 - \ell_3 = 2/v_2 \neq 0$. Choose the quadratic upper entry to realize the determinant of $C(q)$. If $n = 0$ in K, put an upper coupling at position 1 to absorb the upper-band equation. The two couplings end at scalar coordinates. After moving quadratic coordinates first the matrix has form $\left(\begin{smallmatrix} Q & 0 \\ R & aI_r \end{smallmatrix}\right)$ and is similar to $Q \oplus aI_r$, since $Q - aI_2$ is invertible. This follows by conjugating with a lower block unitriangular matrix with lower block $R(Q - aI_2)^{-1}$.

For $s \geq 2$, first suppose $r \neq 2$. Put the first quadratic block on $1, 2$, a scalar coordinate on 3, and the remaining quadratic blocks on successive pairs $4, 5;\ 6, 7; \ldots; 2s, 2s + 1$. The first $2s$ weights are $(1, -1)$ repeated and the other $r - 1$ weights are a nonzero zero-sum tuple

(empty if $r = 1$). The active lower positions inside quadratic blocks are $1, 4, 6, \ldots, 2s$, with w-values $0, -2, -3, \ldots, -s$. Allow one extra lower entry at position 2 into the middle scalar coordinate, with $w_2 = -1$. When $r = 2$, put a scalar at 1, the first quadratic on $2, 3$, a scalar at 4, and the remaining quadratic blocks on $5, 6;\ 7, 8; \ldots; 2s+1, 2s+2$. Use the weight tuple $(1, 1, -2)$ followed by $s - 1$ pairs $(1, -1)$. The internal lower positions are $2, 5, 7, \ldots, 2s + 1$, with w-values $1, -1, -2, \ldots, -(s - 1)$. The extra lower entry at position 3 has $w_3 = 0$. All these values follow from the increments in Lemma 12.6.

In either arrangement, letting t_j be the j-th quadratic lower entry, set the extra lower entry to $-\sum_j t_j$. Then the second lower equation reads

$$t_1 - \sum_{j=2}^{s} (j - 1)t_j = 0.$$

Its nonzero coefficients include $1, -1$, so choose a nonzero zero-sum tuple on the active coefficients, divide by those coefficients and set the other $t_j = 1$. The quadratic blocks now have nonzero lower entries. Choose their diagonal pairs with the given trace, adjusting the first pair if $n = 0$ in K by the nonzero coefficient $1/v_1$ or $2/v_2$. Then set the internal upper entries by the determinant. If $n = 0$ in K, absorb the first upper constraint by adding one upper coupling from the following quadratic block to the scalar center. Its index is 3 or 4 in the two arrangements, respectively. The only case where this coefficient vanishes is the first arrangement in characteristic three with $3 \mid n$, treated below. All off-block couplings go from quadratic blocks to a scalar coordinate, so the same shear proves the required similarity.

If $3 \mid n$ in characteristic three, the dimension and trace equations give $s(\operatorname{tr} C(q) - 2a) = 0$. Every group of three copies of one block is attained by tensoring the scalar I_3 evaluation with that block: tensor the first three inputs with the identity and the fourth with the block. If $\operatorname{tr} C(q) \neq 2a$, both multiplicities are divisible by three. Otherwise remove triples from both types, leaving equal residual counts between zero and two since $r + 2s \equiv 0 \pmod 3$. Each remaining pair $(a) \oplus C(q)$ is a traceless order-three value.

In characteristic two with n odd, r is odd. Remove groups of four quadratic blocks using the I_4 evaluation tensored with one block. We may thus assume $s \leq 3$. If $s = 0$ the residual target is zero by its trace. If $s = 1, r = 1$, use order three. If $s = 1, r \geq 3$, use the arrangement of the first paragraph with all $v_i = 1$ and lower entries $1, 1$ at $2, 3$. If $s = 2, 3$, use the first $s \geq 2$ arrangement with all $v_i = 1$. Its lower equation reduces modulo two and is satisfied by $t_j = 1$ for all j. Since $n \neq 0$ in K, no additional diagonal or upper equation occurs. The determinant choices and scalar shear complete this case. □

Lemma 12.12. *Every traceless target of order $n \geq 3$ with $n = 0$ in K and minimal polynomial degree at most three is an s_4 value. Every traceless target of order five is also an s_4 value over any field.*

Proof. A target whose minimal polynomial is a power of a linear factor is covered by Lemma 12.10. Suppose first that the characteristic p is odd and $p \mid n$. For a split minimal polynomial with two distinct eigenvalues, both primary dimensions are multiples of p, by the total dimension and trace equations. Each primary piece is then covered by Lemma 12.10. With three distinct eigenvalues the target is semisimple. If $n \geq 5$, one eigenvalue has multiplicity at least two. Group two eigenvectors at the other eigenvalues into a cyclic order-two block Q. The $s = 1$ construction of Lemma 12.11 applies with the two exterior scalar coordinates at the repeated eigenvalue. The remaining scalar lines may have any eigenvalues: all the couplings are into the two chosen coordinates and the diagonal conditions use only the prescribed total trace. The smaller order three is already covered. If the minimal polynomial is irreducible quadratic, its multiplicity is divisible by p. If it is irreducible cubic, its multiplicity is divisible by p unless $p = 3$. For $p = 3$, either the multiplicity is divisible by three or

the individual blocks have trace zero. For multiplicity divisible by p, tensor an I_p evaluation from Lemma 12.10 with the companion block as in Lemma 12.11. The remaining cubic blocks use order three. A linear and irreducible quadratic factor is exactly Lemma 12.11 or order three or four.

Now take characteristic two and n even. In the split case, every multiplicity of a distinct eigenvalue is even, since their number is at most three, and two odd multiplicities would give a nonzero trace. If a square linear factor occurs with another linear factor, the same reasoning applies after counting its size-two Jordan blocks separately. Thus the target is a sum of traceless order-two pieces, made of size-two Jordan blocks and pairs of equal scalars. Pair or triple these pieces and use orders four and six. For irreducible quadratic minimal polynomial, if the trace of the companion is nonzero then its multiplicity is even and the blocks pair. If its trace is zero, the blocks group in pairs and triples. For irreducible cubic minimal polynomial, the multiplicity is even and pairs use order six. In the mixed linear and quadratic case, the scalar multiplicity is even. When the quadratic companion is traceless, all the resulting order-two pieces pair or triple. Otherwise the quadratic multiplicity is even and those companions pair into order-four blocks. The scalar part, if of order two, joins one such pair and uses order six, while larger scalar parts use Lemma 12.10.

Finally consider order five. A cyclic target and a target with invariant-factor degrees $(1, 4)$ are already covered by the cyclic and scalar-plus-cyclic constructions. The remaining degree patterns are $(2, 3), (1, 1, 3), (1, 2, 2), (1, 1, 1, 2)$, and five scalar factors. A nonsplit pattern must consist of a linear part and an irreducible quadratic, with multiplicities $1, 2$ or $3, 1$, by invariant-factor divisibility. Use Lemma 12.11. If $5 \neq 0$ in K, a split target in Jordan form uses the first upper band and diagonal of Lemma 12.6 with weights $(1, -1, 1, -1)$. When $5 = 0$ the minimal polynomial degree is at most three and the dividing-characteristic proof applies. □

Lemma 12.13. *A traceless direct sum of at most two cyclic blocks of total order at least three is an s_4 value.*

Proof. The cyclic case, a scalar with a cyclic block, orders three through five, and orders at least eight have been proved. For two blocks the remaining order pairs are $(2, 4)$, $(3, 4)$, $(2, 5)$ and $(3, 3)$. In the first three take the three-dimensional block and triples from Lemma 12.5, using the fourth coordinate of the larger Fiedler block together with the two first coordinates. The outside for $(2, 4)$ is an edge and an isolated vertex. Pair the isolated vertex with one endpoint in a nonzero class and place the other endpoint in the zero class. For $(3, 4)$ the outside consists of two edges. Pair one vertex from each edge in a nonzero class and put the other two in the zero class. For $(2, 5)$ the surviving coordinates $2, 3, 5$ of the larger block form a path, with center 3. Pair the endpoints in one class and the center with the isolated vertex from the smaller block in another. These statements follow from the Fiedler entry formulas.

For $(3, 3)$ the invariant factors of the two cubics f, g are their greatest common divisor and least common multiple, omitting a constant. This follows from the Smith form of $\operatorname{diag}(f, g)$ over $K[x]$. Unless $f = g$, this gives the cyclic case, a scalar plus a cyclic block, or the $(2, 4)$ case. If $f = g$ and $2 \neq 0$ in K, each order-three block has trace zero. If $f = g$ in characteristic two, use order six. □

13 Primary blocks over finite fields

Let K be finite of characteristic p. For a monic irreducible $q \in K[x]$ and $e \geq 1$, the coefficient of $x^{e \deg q - 1}$ in q^e is e times the coefficient of $x^{\deg q - 1}$ in q. Thus the primary cyclic block $C(q^e)$ has trace $e\tau_q$, where $\tau_q = \operatorname{tr} C(q)$.

Lemma 13.1. *Fix q and consider a collection of blocks $C(q^{e_i})$ with $\sum_i e_i = E$ divisible by p. If $E \geq 3$, their direct sum is an s_4 value.*

Proof. Let $F = K[x]/(q)$ and let α be the class of x. Over F, the matrix with blocks $\alpha I_{e_i} + J_{e_i}$, where J_e is the nilpotent Jordan block, is traceless of order E. It is attained by Lemma 12.10. Replace every input matrix over F by its K-linear operator on the same space, using a K-basis of F. This is an algebra embedding. Since finite fields are perfect, $q'(\alpha) \neq 0$. Expanding $q(\alpha I_e + J_e)$ gives $q'(\alpha)J_e$ plus terms of degree at least two in J_e. This is J_e times an invertible polynomial in J_e, since that polynomial has nonzero constant coefficient. Its nilpotence index is therefore e. The K-linear block is annihilated by q^e and its minimal polynomial is exactly q^e, because every monic divisor of q^e is a power of q. This polynomial has degree $e \deg q$, the dimension of the block, so the block is similar over K to $C(q^e)$. □

Proof of Theorem 12.1. The infinite-field assertion follows from Lemmas 12.2, 12.9, 12.8 and 12.12. For a finite field, write the target in primary rational canonical form. A traceless primary block of size at least three is covered by Lemma 12.13. Keep the zero lines and the nonzero traceless blocks of order two for later combination.

Characteristic two. Each nonzero-trace block for a fixed q has odd exponent and the same trace τ_q. Pair these blocks within q. An exponent sum at least four is covered by Lemma 13.1. The other pairs are $C(q) \oplus C(q)$. They are attained if $\deg q \geq 2$, by Lemma 12.13, and otherwise are nonzero scalar blocks of order two, kept as exceptions. After removing the pairs, at most one block remains for each q. Their direct sum is therefore cyclic and has trace zero. If nonempty, its dimension is at least three: one remaining block has nonzero trace, and two lines with opposite nonzero traces in characteristic two would have the same primary polynomial.

The nonzero traceless order-two pieces are either λI_2 with $\lambda \neq 0$, or $J_2(\lambda) = \lambda I_2 + E_{12}$, because an irreducible trace-zero quadratic would be inseparable. If at least two exceptions occur, group them in pairs and triples and use orders four and six. For one exception B, the total dimension guarantees another original piece. We use these absorptions.

- A zero line joins B in order three. With a single traceless primary block C of order at least three, if its primary polynomial is $x - \lambda$, the sum is scalar plus nilpotent. Otherwise $J_2(\lambda) \oplus C$ is cyclic, and $\lambda I_2 \oplus C$ is a scalar line plus the cyclic block $\lambda \oplus C$ with complementary trace.

- With a pair C_1, C_2 for the same primary polynomial q, use scalar plus nilpotent if $q = x - \lambda$. If $q \neq x - \lambda$ and $B = J_2(\lambda)$, use the three-cyclic lemma when both C_i have order at least two. If one C_i is scalar, the other has odd exponent at least three since the scalar pair was itself retained. Merge the other with B by coprimeness and use the scalar-plus-cyclic case. If $q \neq x - \lambda$ and $B = \lambda I_2$, split the two scalar lines and merge one into each C_i. This gives two cyclic summands with combined trace zero, so Lemma 12.13 applies.

- Let R be the cyclic remainder. If $x - \lambda$ is coprime to its polynomial, merge with $J_2(\lambda)$, or merge one scalar line and use the other as in the first case. Otherwise $R = J_e(\lambda) \oplus D$, where e is odd, $\lambda \neq 0$, and D is cyclic and coprime to $x - \lambda$, of trace λ and order at least two. For $B = J_2(\lambda)$ and $e = 1$, merge B with D and use the remaining scalar line. For $e \geq 3$, use the three-cyclic lemma. For $B = \lambda I_2$ and $e \geq 3$, split into $\lambda \oplus J_e(\lambda)$ and $\lambda \oplus D$, using the scalar-plus-nilpotent and cyclic cases. For $B = \lambda I_2$ and $e = 1$, the target is $\lambda I_3 \oplus D$, handled below.

For the last target, scale to $\lambda = 1$ and take the Fiedler representative of D, of order d. Its first coordinate and the three scalar lines give an I_4 block. Use the order-four triple in the padding row of Table 4 for this block. Its cubic value is invertible, and all its cross maps with the three triples in Lemma 12.3 are invertible by the determinants specified in the table. If $d \geq 6$ the remaining $d-1$ coordinates are hollow with bandwidth two, so the partition in Lemma 12.3 works with this extra block. If $d = 2$, use the zero triple on the remaining coordinate. If $d = 3$, use order six. If $d = 4, 5$, the two final padding rows in the table give the entire corresponding Fiedler parameter section with I_3 adjoined.

Odd characteristic. For each q of nonzero block trace, remove minimal nonempty collections of exponents with sum zero modulo p until none remains. A sequence of p residues contains a nonempty zero-sum subsequence, by applying the pigeonhole principle to its $p+1$ partial sums. Thus every removed collection has length at most p, and the remaining multiplicity of any q is at most $p-1$. The removed groups are attained by Lemma 13.1 since their exponent sum is a positive multiple of p.

Let r be the maximum remaining multiplicity and b the total number of remaining blocks, omitting this remainder if $b = 0$. Its trace is zero. Distribute these blocks into layers with no repeated polynomial q in any one layer. Each layer is cyclic by coprimeness. When $r \leq 2$, use at most two layers and apply Lemma 12.13 if their total dimension is at least three. A one-dimensional remainder is impossible. A two-dimensional remainder consists of opposite nonzero scalars and is kept as an additional traceless cyclic order-two piece. When $r \geq 3$, put m equal to the least even integer at least r, so $4 \leq m \leq p-1$. There are at least m blocks: otherwise r is odd, $b = r$, and all remaining blocks belong to one q. Their total trace zero would give another zero-sum group. Use m nonempty layers, assigning each q's blocks to distinct layers and filling empty layers first.

We describe a source for these m layers, needed only for $p \geq 5$. In any order $m \geq 4$ invertible in K, choose the weights v_i by concatenating the zero-sum words $(1, 1, -2)$, $(1, 1, 1, -3)$, $(1, 1, 1, 1, -4)$ to length $m-1$. Every integer at least three is a sum of $3, 4, 5$. The weights and the adjacent sums, including the two sums with the endpoint convention $v_0 = v_m = 0$, are nonzero. Form J, C, V as in Lemma 12.6. For the core triple take $A = I_m + J$, $B = J^2$, C. It obeys $B = A^2 - 2A + I_m$, and its cubic value $s_3(A, B, C) = AVA - V = JVJ + JV + VJ$ is triangular with diagonal $v_{i-1} + v_i$. It is invertible. The local fourth-slot image contains the traceless diagonal and first upper band by Lemma 12.6 since the fourth-slot map is the negative of the map L in that lemma.

For the remaining coordinates construct three triples of order two with the same relation $B = A^2 - 2A + I_2$. Choose a nonsquare $d \in K$, put $H = \left(\begin{smallmatrix} 0 & d \\ 1 & 0 \end{smallmatrix}\right)$, and choose three distinct $a \in K$ with $a^2 \neq d+1$, possible since $|K| \geq 5$. For each choose a different nonzero $\delta \in K$ and find $u, w \in K$ with $u^2 - dw^2 = \delta$. The sets $\{u^2\}$ and $\{\delta + dw^2\}$ each have size $(|K|+1)/2$, so they meet. Set

$$A = aI_2 + H, \qquad V = \begin{pmatrix} u & -dw \\ w & -u \end{pmatrix}, \qquad C = \frac{HV}{2d}.$$

Multiplication gives $H^2 = dI_2$, $HV = -VH$ and $V^2 = \delta I_2$. Hence $[A, C] = (H^2V - HVH)/(2d) = V$, and $AVA = (a^2 - d)V$. By (15), the cubic value is $(a^2 - d - 1)V$, which is invertible. The characteristic polynomials of these three A's are distinct irreducible quadratics, and that of the core A is a power of $x - 1$. The characteristic polynomials of these three V's are the distinct $x^2 - \delta$, and the core V is nilpotent.

Apply (15) to a block diagonal triple with the common relation on A, B. Writing A_i, V_i and A_j, V_j for the matrices on the respective row and column blocks, the rectangular map is the composition of $X \mapsto A_iX - XA_j$, $X \mapsto V_iX - XV_j$, and $X \mapsto A_iX - XA_j$, in that order. For each map, coprime characteristic polynomials imply invertibility: an intertwiner

X is annihilated on the right by the characteristic polynomial of the left matrix evaluated on the right matrix, which is invertible by Bézout's identity. Thus all cross maps between different types are invertible, including the core maps. Cross maps with the zero triple are invertible by the cubic values.

This realizes a traceless diagonal and first upper band on the core, with arbitrary cross entries to an outside part having the four-class support of Lemma 12.3. Outside coordinates may be any hollow support of bandwidth two of order at least five. The following smaller supports also admit that partition: an independent set, an edge with one or two isolated vertices, two disjoint edges, or a three-vertex path and one isolated vertex. For a path, pair the endpoints and pair the middle with the isolated vertex. For an edge, pair one endpoint with an isolated vertex and leave the other endpoint in the zero class, or use two such pairs. For two disjoint edges, one endpoint of each makes an independent pair and the other endpoints form the zero class.

Take the first coordinate of each cyclic layer in Fiedler form for the core. Its diagonal is traceless. The remaining coordinates of each layer, called its tail, give a hollow block of bandwidth two. If its dimension h is at least five, the source above applies directly. For $h \le 4$ the only possible tail patterns beyond the small supports just listed are a single tail of length two, three or four, possibly with one or two isolated vertices, within that dimension bound. An edge with isolates already works. For a layer of order four or five (tail length three or four), move its third and fourth coordinates to the core. The only possible additional core entries in that layer are $(1,3)$ and $(3,4)$, which are on the first upper band after ordering these three coordinates consecutively. The new core size is even, between six and $p+1$, and therefore invertible. The surviving coordinates of that tail are 2 or $2,5$, independent, and any other outside coordinate is isolated. The remaining case is a single edge from a cyclic cubic and no other outside vertex. Move coordinate 3 of that layer into the core, leaving one isolated outside vertex and a single new first upper entry. The new size is invertible unless it equals p. In that event the whole remainder consists of a cyclic cubic and scalar lines of total size $p+1$.

The last configuration, more generally one cyclic cubic and scalar lines in even total size $n \ge 6$ invertible in K, is obtained as follows. Put the cubic on coordinates $2,3,4$ and choose nonzero weights with $v_1+v_2=0$ and $v_3+\cdots+v_{n-1}=0$, using pairs and a triple as necessary. The first lower entries at positions $2,3$ can be $t,-t$, with $t \ne 0$, since $w_2 = w_3$. Lemma 12.7 represents the cubic using just those lower entries, diagonal and first upper entries. All these bands are in the weighted image, giving the configuration with any prescribed scalar lines.

Absorption in odd characteristic. Every nonzero traceless order-two matrix in odd characteristic is cyclic. Pair or triple the retained pieces if at least two occur, using order four or the three-cyclic lemma. Let B be a sole retained piece. It joins a zero line by order three, or a traceless primary block by Lemma 12.13. If it joins a remainder with at most two cyclic layers, one layer uses the two-cyclic lemma and two layers of size at least two use the three-cyclic lemma. If one of two layers is scalar, the scalar-plus-cyclic hollowization of Lemma 12.4 concatenated with the hollow companion of B gives a hollow target of bandwidth two. The order-four case is already covered.

If it joins a remainder with $m \ge 4$ layers in characteristic $p \ge 5$, use their original first Fiedler coordinates for the core, and append the two hollow coordinates of B outside. If the original outside has size at least three, apply the bandwidth-two source. For original sizes one and two, the original outside has either only isolated coordinates or a single edge on two coordinates. Adjoining the two coordinates of B gives an edge with isolated vertices or two disjoint edges. For size zero the original remainder consists of m scalar lines. The total size $m+2$ is even and at most $p+1$, hence invertible. If B splits, use a Jordan form and the diagonal and first upper bands of Lemma 12.6. Otherwise merge B with a scalar line into a

cyclic cubic and use the cubic and scalar construction above.

It remains to absorb B into a removed group of ℓ blocks for one primary polynomial q. Write $f = \deg q$ and E for the sum of exponents. For $\ell = 2$, use the three-cyclic lemma if both blocks have order at least two. If one is scalar, its sum with the other block has a hollow bandwidth-two representative, which concatenates with the hollow companion of B. The order is at least five. The case $\ell = 1$ was covered by the single traceless primary block case.

For $p = 3$ the only remaining group has three blocks with exponents all 1 or all 2 modulo three. If q is coprime to the polynomial of B, merge B with one of the three, choosing a scalar block if any. The resulting three cyclic blocks either have sizes all at least two or total size at least eight, except when all three original blocks are scalar and the total size is five. Use the three-cyclic lemma or the order-five result. If q is not coprime, it must be $x - \mu$ with $\mu \neq 0$ and $B = (\mu) \oplus (-\mu)$, since $\tau_q \neq 0$. If the exponents are 2 modulo three, put (μ) with one block, giving a traceless scalar plus nilpotent. Combine $(-\mu)$ with a second block by coprimeness and use the two-cyclic lemma on that block and the third, of total trace zero. If the exponents are 1 modulo three and one is at least four, join $(-\mu)$ to that block, making a traceless cyclic block, and put (μ) with the other two, making a traceless scalar plus nilpotent. If all three exponents are one, use order five.

For $p \geq 5$ and even $\ell \geq 4$, use the ℓ primary Fiedler roots as the core and keep B outside. The core has invertible size, and the outside size is $h = fE - \ell + 2$. When $h \geq 5$ the bandwidth-two source applies. When $h \leq 4$, we have $fE \leq \ell + 2 \leq p + 1$. Since E is a positive multiple of p, this gives $f = 1, E = p$. Now $\ell \geq p - 2$, so evenness gives $\ell = p - 1$. Thus one exponent is two and the others one. The outside is B's edge and an isolated vertex. For odd $\ell \geq 3$, add the first hollow coordinate of B to the ℓ primary roots. The core size $\ell + 1$ is even and at most $p + 1$, so is invertible. The outside size is $h = fE - \ell + 1$. If $h \geq 5$ the bandwidth-two source applies. If $h \leq 4$, then $fE \leq \ell + 3 \leq p + 3 < 2p$, so $f = 1, E = p$. Since p and ℓ are odd, $h = p - \ell + 1$ is one or three. When $h = 1$, all exponents are one. When $h = 3$, one exponent is three or two are two, leaving an edge and an isolated vertex, or an independent set. These are precisely the small supports covered above.

Any zero lines not used in an absorption are attained by zero inputs. Taking the four block diagonal sums of the inputs for the pieces produces the block diagonal representative in one evaluation. Simultaneous similarity gives the target. □

14 The quartic classification

Proof of Theorem 1.2. Theorem 1.3 gives the cubic classification. For a quartic in size one, every word is the same scalar product and the image is $\{0\}$ or K according as the coefficient sum is zero or nonzero. Theorem 6.2 treats size two. Assume $n \geq 3$. Lemma 1.1 identifies vanishing of the evaluated trace with vanishing of the formal trace and excludes nonzero quartic identities in this range.

The zero polynomial has zero image. Suppose $p \neq 0$ has zero formal trace. Deleting one fixed letter from two cyclically equivalent quartic words gives cyclically equivalent cubic words. Thus deleting that letter defines a linear map on the spaces of cyclic classes. The formal trace of the cubic obtained by substituting the identity is the image of the quartic formal trace under this map, so it is zero. If one such cubic is nonzero, its image is $\mathfrak{sl}_n(K)$ by Theorem 1.3, and each value is a value of p. Otherwise p is proper. If a one-variable coefficient from Lemma 8.1 is nonzero, Proposition 8.4 applies. If all four vanish, Lemma 8.1 puts p in the span of the six commutator products. Proposition 8.5 applies except to scalar multiples of s_4, for which Theorem 12.1 gives the result. Cyclicity of trace supplies the reverse inclusion in every case.

Suppose p has nonzero formal trace. If one substitution of the identity gives a cubic of nonzero formal trace, Theorem 1.3 gives full image. If p is proper, Corollary 8.6 applies. Otherwise all single-identity substitutions have zero formal trace and Lemma 10.1 supplies the order-three scalar values. For $|K| \geq 3$, Lemma 10.2 gives auxiliary states for Proposition 9.1. Over $\mathbf{F}_2$ use Proposition 11.1. In either case the image is $M_n(K)$. □

A Coefficient and matrix certificates

A.1 Formal auxiliary coefficients

Order the 24 words lexicographically. Each coefficient condition below is a row of length 24 over the prime field, applied to the column of word coefficients. For each identity substitution and each of the two cyclic classes of the resulting three-letter words, put a 1 in a column if deleting the specified letter gives a word in the specified class, and 0 otherwise. The eight rows so obtained form the initial list. A cyclic coefficient row for the original quartic has a 1 precisely at the four words in its cyclic class. A properness row for a deleted letter and a three-letter word has a 1 precisely where the deletion gives that word itself.

For each of the six unordered pairs, orient the pair in increasing slot order as X, Y. The linear forms that are the coefficients of q, h are obtained directly by the rule defining H, G in Section 9. For the swapped orientation denote the resulting forms by $\widetilde{q}, \widetilde{h}$. Failure to find a pair with $\kappa \neq 0$ and both boundary polynomials formally nonzero in at least one orientation is an intersection of six unions. For one pair the union has five branches: impose $\kappa = 0$, or impose vanishing of one of q, h and one of $\widetilde{q}, \widetilde{h}$. Vanishing of a polynomial here means appending the rows of all its coefficients. The value of κ is unchanged by swapping the pair, by (12).

To reduce a list, scan the columns from left to right. Use the first remaining row with nonzero entry in the column as pivot, scale it to give entry one and subtract its multiples from all other rows. For pruning, sort the resulting distinct lists by row rank, and discard a list if its span contains the span of a previously retained list. Its null space is then already covered. To test a containment, cancel each candidate row in the pivot columns using the rows of the proposed larger space and check that the remainder is zero. In detail:

1. Start with the reduced initial list as the sole retained space.

2. At the next pair, form five new row lists from every retained space. Append either the κ row or the coefficient rows for one of $(q, \widetilde{q}), (q, \widetilde{h}), (h, \widetilde{q}), (h, \widetilde{h})$. For any of these polynomials use the four monomials $u_1u_2, u_1v_2, v_1u_2, v_1v_2$, collecting one row per monomial by the defining word rule. For the κ row use a 1 exactly when Y immediately follows X in the cyclic word.

3. Reduce the lists and keep one copy of each reduced list. Discard each list containing a previously retained span in rank order. Repeat step 2 for the next pair.

4. After the last pair, reduce each original cyclic coefficient row and each properness row against each surviving list. Check whether all remainders of the first kind are zero or all remainders of the second kind are zero.

The intermediate results, in pair order $(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)$, are

	12	13	14	23	24	34
number of row spaces	5	25	51	18	2	2
ranks	5, 11	6, 12, 15	6, 16	6, 16, 18	6, 18	6, 18.

The final rank-six space of rows contains the cyclic coefficient rows. The rank-eighteen space contains the properness rows. In the rational calculation all nonzero pivot numerators are powers of two up to sign. Inductively all denominators are powers of two. Consequently the row operations take place over $\mathbb{Z}[1/2]$. Each row containment used to prune a branch, and each final containment, is checked by subtracting multiples of echelon rows with leading entry one. Such equalities survive reduction in every odd characteristic, even when new containments occur. They also hold over every field of characteristic zero. In characteristic two the same table results by separate row reduction over $\mathbf{F}_2$. The supplied program `coefficient_two_color_cover.py` constructs the rows from the words and carries out the calculation, including the final containment tests. Its argument is 0 for the rational calculation and 2 for the characteristic-two calculation.

A.2 Binary coefficients

Use the same word ordering with indices starting at zero. For each ordered pair and each of the four choices of $v \in \mathbf{F}_2^2$ use $u_i = 1 + v_i$ and evaluate the conditions in Proposition 9.1. The entries r, s, q, h and the two coefficients in the last condition can be computed on size-two matrix units. In particular, r is the $(1,1)$ entry at $X = E_{12}, Y = E_{21}$, s is that entry with the two inputs exchanged, q is the $(1,2)$ entry at $X = E_{12}, Y = E_{11}$, and h is the $(2,1)$ entry at $X = E_{21}, Y = E_{11}$. Interchanging the diagonal states computes q', h'. The trace at $X = E_{12}, Y = E_{21}$ gives t. The coefficients $H(v,v,u), G(v,u,u)$ in the third condition are the $(2,1)$ entries at $Y = E_{21}$ with $X = E_{22}, E_{11}$, respectively. In each evaluation the auxiliary inputs are diagonal with entries u_i, v_i. For the reversed condition reverse the letters of each word before evaluating the same matrix-unit tests. This gives the reversed test of Proposition 9.1.

The complete calculation consists of the following steps.

1. Row reduce the single-identity cyclic equations over $\mathbf{F}_2$. Form the 2^{20} elements of the kernel by taking each sum of a basis subset. Discard any mask on which all six original cyclic coefficient rows vanish, and any mask on which all properness rows vanish. There are 786048 masks left.

2. For each ordered pair and each state, evaluate the three conditions and discard every mask that satisfies one of them with $t \neq 0$. This leaves 1272 masks.

3. For each representative below, apply each of the 24 slot permutations to its words, with or without reversing their order, and with or without complementing all 24 bits. Remove duplicates within each resulting set. Check membership of each resulting mask among the masks from step 2, mark it and check that no other representative marks it. The 24 sets mark all 1272 remaining masks.

30	221	783	1495	2600	3900
8587	9417	10252	12767	13085	14220
41870	63000	64755	275151	325532	377795
832207	832444	837351	838668	846652	923760

The supplied `independent_coefficient_cover.py` performs these steps by evaluating the 24 original words for each test. All bit operations are operations over $\mathbf{F}_2$. The set equality in step 3 is checked directly.

A.3 Linear maps for the binary representatives

The following table specifies the repeated slots and the free edge slot for the six representatives with an $[A^2, B]$ factor. The fourth slot not designated as repeated or free is the

slot for B. The identities follow by expanding the original words after identifying the two repeated slots.

mask	repeated	free	output
30	$2,3$	1	$X[A^2,B]$
221	$3,4$	1	$X[A^2,B]$
8587	$2,3$	4	$[A^2,B]X$
14220	$1,3$	2	$X[A^2,B]$
41870	$1,3$	4	$[A^2,B]X$
325532	$1,2$	3	$[A^2,B]X$

The repeated-slot formulas for $783, 2600, 832207$ in Section 11 are checked by the same expansion.

The tables below give the fixed inputs for the square and rectangular maps used in Section 11. For a row, recover the words of the polynomial by reading the 24 bits of its mask. Decode each fixed matrix and put it into the indicated slot. For two sizes, take the direct sums of the fixed matrices in each slot. Form a column of the free-input operator by putting a matrix unit in the free slot and summing the original words with bit one. For a rectangular operator, put the unit in the indicated rectangular block and set all other free blocks to zero. Write the entries of the resulting block in row order.

In Table 1 the combined map on order three has rank eight and every output has bottom-right entry zero. The maps on the other three blocks act separately, so their ranks sum to eight and each has full rank. In Table 2 both rectangular operators have rank six. In the two square blocks, adjoin as extra columns the three specified triangular basis elements of order two and the four specified basis elements of order three, respectively. In each case the augmented matrix has the same rank as the original. In Table 3 the square maps have the listed ranks. In each rank-three row the decoded annihilator has rank and trace one, and its trace pairing with every output column is zero. The supplied verification program also checks preimages for the paired records.

A.4 Fixed input tuples

The tables encode each matrix by the integer obtained by reading its entries as bits in row order. The top-left entry is the coefficient of 2^0. The free position is indexed from 1, and each fixed tuple lists the other three positions in increasing order. In Table 2, C refers to the first size-three space of Section 11, and K to the second. To check those rows without preimages, augment the square operators by the indicated target columns and compare ranks. The rectangular operators have rank six. Table 3 encodes the annihilator matrix itself, for the trace pairing.

Table 1: Size-two sources with a scalar border.

Mask	Free position	Size-two tuple	Scalar tuple
9417	2	$(1,5,6)$	$(0,1,1)$
12767	1	$(1,6,3)$	$(1,1,1)$
13085	2	$(2,7,4)$	$(0,0,1)$
63000	3	$(1,6,3)$	$(0,0,0)$
64755	3	$(1,6,3)$	$(0,0,0)$
377795	1	$(1,6,11)$	$(1,1,1)$
832444	2	$(1,3,5)$	$(0,1,0)$
837351	2	$(1,6,12)$	$(0,1,1)$
838668	2	$(1,3,5)$	$(0,1,1)$

Mask	Free position	Size-two tuple	Scalar tuple
846652	2	$(1,3,7)$	$(0,0,0)$
15930563	4	$(1,3,6)$	$(0,0,0)$
15938547	1	$(1,3,5)$	$(0,0,1)$
15939864	2	$(1,3,13)$	$(0,0,1)$
15944771	4	$(1,3,5)$	$(0,0,1)$
16399420	2	$(1,3,7)$	$(1,0,1)$
16502064	1	$(1,3,5)$	$(0,1,1)$
16712460	4	$(1,6,3)$	$(0,0,1)$
16714215	4	$(1,6,3)$	$(0,0,1)$
16764130	3	$(1,6,3)$	$(0,0,0)$
16764448	4	$(1,3,6)$	$(1,0,1)$
16767798	2	$(1,5,6)$	$(0,1,0)$

Table 2: Size-two sources with size-three cores.

Mask	Free position	Core type	Size-two tuple	Size-three tuple
1495	1	K	$(1,8,7)$	$(415,62,294)$
1495	1	C	$(2,4,3)$	$(229,475,274)$
3900	1	K	$(1,15,14)$	$(143,294,415)$
3900	1	C	$(1,15,14)$	$(143,294,415)$
10252	1	K	$(4,14,3)$	$(58,214,293)$
10252	1	C	$(4,14,3)$	$(443,507,201)$
275151	1	K	$(1,3,5)$	$(134,323,299)$
275151	1	C	$(1,3,5)$	$(134,323,299)$
832207	1	K	$(1,10,7)$	$(160,117,20)$
832207	1	C	$(1,10,7)$	$(160,117,20)$
923760	1	K	$(2,8,12)$	$(313,220,143)$
923760	1	C	$(1,6,8)$	$(481,1,354)$
15853455	1	K	$(1,6,8)$	$(24,79,39)$
15853455	1	C	$(1,6,8)$	$(24,79,39)$
15945008	1	K	$(1,3,7)$	$(8,191,464)$
15945008	1	C	$(1,2,5)$	$(327,250,284)$
16766963	1	K	$(1,1,14)$	$(64,46,459)$
16766963	1	C	$(7,2,1)$	$(105,69,250)$
16773315	1	K	$(1,1,7)$	$(445,407,415)$
16773315	1	C	$(2,5,8)$	$(133,392,442)$
16775720	1	K	$(1,6,14)$	$(323,46,333)$
16775720	1	C	$(1,14,8)$	$(500,181,481)$

Table 3: Square operators in the even-order construction.

Mask	Free position	Size-two tuple	Rank	Annihilator
1495	1	$(1,1,14)$	3	8
3900	1	$(1,5,11)$	3	10
9417	2	$(1,5,3)$	4	–

Mask	Free position	Size-two tuple	Rank	Annihilator
10252	3	$(1, 1, 7)$	3	8
12767	4	$(1, 3, 6)$	4	–
13085	1	$(1, 5, 7)$	4	–
63000	3	$(1, 6, 3)$	4	–
64755	3	$(1, 6, 3)$	4	–
275151	1	$(1, 3, 5)$	4	–
377795	2	$(1, 3, 5)$	4	–
832444	1	$(1, 3, 5)$	4	–
837351	1	$(1, 3, 13)$	4	–
838668	1	$(1, 3, 5)$	4	–
846652	1	$(1, 3, 5)$	4	–
923760	2	$(1, 1, 14)$	3	1

B Products with a common commutator factor

Lemma B.1. *Let C, D be endomorphisms of an n-dimensional vector space over K, with I, C, D independent. Assume $n \geq 3$ if $|K| > 2$, and $n \geq 5$ if $K = \mathbf{F}_2$. If v, Cv, Dv are dependent for every v, then there is a line Kw and scalars λ, μ such that both $C - \lambda I$ and $D - \mu I$ have image in Kw.*

Proof. First take $|K| > 2$. Choose x with x, Cx independent and replace D by $D - aI - bC$ so that $Dx = 0$. Put $W = \operatorname{span}\{x, Cx\}$. For any y,

$$(x + ty) \wedge (Cx + tCy) \wedge tDy = 0 \qquad (t \in K).$$

The coefficient of t^3 vanishes by the hypothesis at y. Dividing by t at the nonzero t gives a linear polynomial vanishing at two distinct points. Its constant term gives $Dy \in W$. Write $Dy = \alpha(y)x + \beta(y)Cx$ with linear functionals α, β. The coefficient of t after division is $-x \wedge Cx \wedge (\alpha(y)y + \beta(y)Cy)$, so $\alpha(y)y + \beta(y)Cy \in W$. If D has rank two then α, β are independent. The vectors of $\ker\beta$ outside $\ker\alpha$ span $\ker\beta$: take one such vector and add each vector in $\ker\alpha \cap \ker\beta$ to it. The inclusion then gives $\ker\beta \subset W$. This forces $n = 3$ and $\ker\beta = W$. The displayed inclusion implies $\beta(Cy) = -\alpha(y)$ wherever $\beta(y) \neq 0$, and hence everywhere, because the vectors outside a hyperplane span the space. As $\alpha(x) = 0$ and α, β are independent, $\alpha(Cx) \neq 0$. Thus $D(Cx)$ is a nonzero multiple of x and $C^2x \notin W$, contradicting the hypothesis at Cx.

We have shown that D has rank one, say $Dy = \phi(y)w$. For $\phi(y) \neq 0$, $y \wedge Cy \wedge w = 0$. If $\phi(y) = 0$, choose z with $\phi(z) \neq 0$ and apply the identity at $y + tz$ for $t \neq 0$. It is linear in t because its quadratic term vanishes at z. It therefore holds at $t = 0$ as well. Subtracting the identity at y from that at $y + w$ gives $y \wedge Cw \wedge w = 0$ for all y. If $Cw \notin Kw$, a vector outside $\operatorname{span}\{w, Cw\}$ contradicts this identity. The induced map on the quotient by Kw preserves every line and hence is scalar, by applying it to a basis and to pairwise sums. Thus $C - \lambda I$ has image in Kw. Undoing the replacement of D gives the assertion for the original pair.

Now let $K = \mathbf{F}_2$ and $n \geq 5$. Apart from I, the six nonzero members of $\operatorname{span}\{I, C, D\}$ form three pairs of the form $A, A + I$. Their kernels cover all nonzero vectors by the dependence hypothesis. Within each pair the kernels are disjoint except at zero. If all six maps have rank at least two, the two kernel dimensions in a pair are each at most $n - 2$ and have sum at most n. Therefore their union contains at most $2^{n-2} + 2$ nonzero vectors. This bound follows by maximizing the sum of the two powers of two with the stated constraints, by transferring units to the larger dimension. Three pairs cover at most $3(2^{n-2} + 2) < 2^n - 1$

nonzero vectors. Thus a rank-one map is present. Replace C, D by two members giving a basis with I, with the new D of rank one. It suffices to prove that C is scalar modulo the image line Kw of D, since the conclusion is preserved under a change of basis of this three-dimensional operator space. Write $Dy = \phi(y)w$.

If $\phi(w) = 0$, apply dependence at y and $y+w$ for every y with $\phi(y) = 1$. Modulo Kw, the images of these two vectors under C lie in the same line $K\bar{y}$. Hence the image of Cw in the quotient is in the intersection of at least two distinct such lines and is zero. The induced C on the quotient maps each vector in the affine hyperplane $\phi = 1$ either to itself or to zero. If two vectors there have different choices, take a third independent of them in the same affine hyperplane. The sum of the three vectors belongs to the hyperplane, but its image cannot be either zero or the whole sum, a contradiction. The hyperplane spans the quotient, so the induced map is scalar.

If $\phi(w) = 1$, identify the quotient by Kw with $\ker\phi$. Write B for the linear map obtained by restricting C to $\ker\phi$ and projecting back along Kw, and c for the projection of Cw. For each nonzero $u \in \ker\phi$, applying dependence at $u + w$ gives $Bu + c \in Ku$. For independent u, v add the relations for $u, v, u + v$ to obtain $c \in \operatorname{span}\{u, v\}$. Intersecting these spans gives $c = 0$, since $\dim\ker\phi \geq 4$. Hence B preserves all lines and is scalar. This proves the claim in either case. □

We use the trace-moment criterion (2) for a cyclic commutator factor.

Lemma B.2. *Let K be any field and $n \geq 2$. For every pair of traceless matrices C, D of order n, there is a cyclic matrix A for which C, D both lie in the image of $X \mapsto [A, X]$.*

Proof. We first treat order two. The subspace of matrices orthogonal to C, D has dimension at least two and contains the identity. Every nonscalar matrix A in it is cyclic, and the two moments for C, D vanish.

We will also use a nilpotent choice for a single traceless matrix T of order $m \geq 3$ when $|K| > 2$. Write T in a direct sum of column companion matrices of its invariant factors. Let h be the number of its nonzero first-subdiagonal entries. If $h \geq 2$, choose h nonzero scalars summing to zero. For h even use pairs $(1, -1)$, and for h odd use one triple $(1, u, -1-u)$ with $u \notin \{0, -1\}$ and then pairs. Diagonal conjugation makes the subdiagonal entries equal these scalars because the entries lie on disjoint paths within the companion blocks. The resulting matrix is upper Hessenberg and has zero sum on the first subdiagonal, so its trace pairing with every power of the upper nilpotent shift $J_m = E_{12} + \cdots + E_{m-1,m}$ vanishes. If $h = 0$, all invariant factors are linear and T is triangularizable. If $h = 1$, there is a single quadratic invariant factor and at least one linear factor dividing it, so T is again triangularizable. In either case take T upper triangular, and all the moments with J_m vanish by tracelessness. Conjugating J_m back gives the desired regular nilpotent matrix.

Suppose next that $|K| \geq n$ and $|K| > 2$, allowing K infinite. If C, D act on a residual space W of dimension at least three and its identity does not lie in their span, there is a vector $v \notin \operatorname{span}(Cv, Dv)$. If both matrices are zero, take any nonzero v. When their span is one-dimensional, the contrary assumption for a generator C makes C invertible and C^{-1} line-preserving, hence C scalar. When I, C, D are independent, choose a vector witnessing independence of v, Cv, Dv if there is one. Otherwise Lemma B.1 gives a line Ku and independent row functionals ρ, σ such that $C = \lambda I + u\rho$, $D = \mu I + u\sigma$. If $\lambda = \mu = 0$, take v outside Ku. Otherwise choose v with $\rho(v) = \lambda$, $\sigma(v) = \mu$, outside Ku. If the affine solution set lay in Ku, its translation space, of dimension at least one, would equal Ku, and then every vector in the solution set would have zero values under ρ, σ. Now $Cv = \lambda(v + u)$ and $Dv = \mu(v + u)$, so v is outside their span. Choose a functional ϕ with $\phi(v) = 1$ and $\phi(Cv) = \phi(Dv) = 0$. In the decomposition $Kv \oplus \ker\phi$, both first diagonal entries are zero, and the two lower-right compressions are traceless.

Repeat until a residual block has order two or the identity lies in the span of its two compressions. In order two use the preceding argument. In the identity-span case its order is zero in K, by taking traces, and the residual span is generated by the identity and at most one other matrix. Use the preceding nilpotent construction on that matrix. The identity is orthogonal to every power of a regular nilpotent in that order. Thus in both cases we have a cyclic A_0 for the final block. Adjoin distinct scalar blocks with scalars outside the spectrum of A_0 in K, one for every removed entry. There are at most as many excluded scalars as the order of A_0, so there are enough scalars. The resulting A is cyclic and its centralizer is the direct sum of the block centralizers because the block minimal polynomials are coprime. The removed entries of C, D were zero, so both matrices are orthogonal to this centralizer.

It remains to treat finite fields of order $q < n$. When C, D are dependent, the single-target commutator construction in Lemma 2.1 gives a cyclic factor for a nonscalar generator, and a regular nilpotent works for a scalar generator because it is traceless. When $q > 2$ and the identity lies in their span, the same nilpotent argument used for the residual block above applies.

Now suppose I, C, D are independent. We may assume $n \geq 4$ if $q > 2$ and $n \geq 5$ if $q = 2$, since the small binary endpoints are checked below. If v, Cv, Dv are dependent for every $v \in K^n$, Lemma B.1 gives scalars λ, μ and a common image line Ku for $C - \lambda I, D - \mu I$. Choose a regular nilpotent N with kernel Ku. The moments against N^j, $j \geq 1$, vanish for both matrices, since N^j kills u and has trace zero. The zeroth moments vanish by hypothesis.

Suppose therefore that v, Cv, Dv are independent for some v. Put $U = \operatorname{span}(Cv, Dv)$ and choose a complement W containing v, of dimension $m = n - 2$. In the basis (Cv, Dv) of U and a basis of W starting with v, let F, G be the upper-right 2-by-m blocks of C, D. Their first columns are $(1, 0)^{\mathsf{T}}, (0, 1)^{\mathsf{T}}$. We identify W^* with the field L of order q^m so that the rows of F, G, represented by f_1, f_2 and g_1, g_2, satisfy $f_1 g_2 - f_2 g_1 \neq 0$. To do so, choose independent field images $\beta_2, \ldots, \beta_m$ of the last $m - 1$ dual basis vectors. In the image β_1 of the first vector, the determinant is a monic quadratic by the first-column conditions. Since $q^m - q^{m-1} > 2$ for $m \geq 3$ or for $m = 2, q > 2$, we can choose β_1 outside the span of the other images and outside the two possible roots.

Let P act on W so that its dual action is multiplication by an element $\eta \in L$ of degree m over K. Let $Z : U \to W$ be a lower shear parameter, with columns represented by $z_1, z_2 \in L$ using the field trace pairing. The lower-right compressions after conjugation by $\left(\begin{smallmatrix} I & 0 \\ Z & I \end{smallmatrix}\right)^{-1}$ are $C_{22} - ZF$ and $D_{22} - ZG$. Their moments against P^j, $0 \leq j < m$, vanish precisely when

$$\operatorname{Tr}_{L/K}(\eta^j (f_1 z_1 + f_2 z_2)) = \operatorname{tr}(P^j C_{22}),$$
$$\operatorname{Tr}_{L/K}(\eta^j (g_1 z_1 + g_2 z_2)) = \operatorname{tr}(P^j D_{22}).$$

Here $\operatorname{Tr}_{L/K}$ is the field trace, the trace of multiplication on L over K. The left formulas follow by expressing ZF and ZG as sums of column-row products. The field trace pairing is nondegenerate and the powers of η form a basis. The two equations therefore prescribe $f_1 z_1 + f_2 z_2$ and $g_1 z_1 + g_2 z_2$ uniquely, so the chosen determinant gives a solution. The upper-left compressions are now traceless as well. Apply the order-two argument to choose a cyclic B for them. Its characteristic polynomial is coprime to that of P if $m \geq 3$. If $m = 2$, choose a shift $B + cI_2$ whose determinant differs from that of P, avoiding at most two values of c because the determinant is monic quadratic in c. There are enough choices since $q > 2$. The shift does not change the centralizer, hence preserves cyclicity and the orthogonality to that centralizer. Then $\operatorname{diag}(B, P)$ is cyclic and the moments vanish by its block centralizer. Undo the shear to obtain A.

For the remaining identity-span case over $\mathbf{F}_2$, the order is even. It suffices to handle I, C with C nonscalar and traceless. Write $n = 2 + m$ with $m = 2e$, and choose a decomposition of these dimensions with the second row f_2 of the top-right block F of C nonzero. This is

possible by taking w, Cw independent, putting w in the second summand while Cw has a nonzero projection to the first, and choosing the basis of the first summand. Choose a monic irreducible polynomial g of degree e with $g(0) \neq 0$, and set $L = \mathbf{F}_2[t]/(g^2)$. Let ℓ take the coefficient of t^{m-1} in the representative of degree less than m. The pairing $(u, v) \mapsto \ell(uv)$ on L is nondegenerate because multiplying any nonzero representative of degree i by t^{m-1-i} gives a nonzero coefficient in degree $m-1$. Identify the dual of the second summand with L, mapping f_2 to 1, and write $f_1 \in L$ for the other row. Let P act with dual action multiplication by t. Every polynomial in P has trace zero, because every root of the separable polynomial g has algebraic multiplicity two in P. Also P is cyclic and its minimal polynomial is coprime to t^2.

Represent the two columns of a lower shear Z by $z_1, z_2 \in L$ using the pairing $\ell(uv)$. The equations for the lower-right compression of C are

$$\ell(t^j(f_1 z_1 + z_2)) = \operatorname{tr}(P^j C_{22}), \qquad 0 \leq j < m.$$

They specify $f_1 z_1 + z_2$ uniquely. The lower-left entry of the upper-left compression $C_{11} + FZ$ is zero if $\ell(z_1) = (C_{11})_{21}$. Choose $z_1 = (C_{11})_{21} t^{m-1}$ and solve for z_2. The upper-left compression is traceless and has zero lower-left entry, so it is orthogonal to the centralizer of $J_2 = E_{12}$. The lower-right compression is orthogonal to $K[P]$. The cyclic matrix $\operatorname{diag}(J_2, P)$ has a traceless centralizer, and therefore satisfies the moments for I as well. Undo the shear.

For orders three and four over $\mathbf{F}_2$, the finite check following the proof verifies (2) simultaneously for every ordered pair. All fields and orders are thus covered. □

B.1 Binary common-factor check

The finite verification uses the following exact algorithm for each $n = 3, 4$.

1. Enumerate the matrices A over $\mathbf{F}_2$ whose powers $I, A, \ldots, A^{n-1}$ are independent. Store the lists of linear forms $T \mapsto \operatorname{tr}(A^j T)$, $1 \leq j < n$, on the traceless matrices. Independence is equivalent to cyclicity by the rational canonical form.

2. Enumerate all traceless C up to conjugation by the elementary transvections $I + E_{ij}$, $i \neq j$. These generate $\mathrm{GL}_n(\mathbf{F}_2)$ by row reduction. For each representative, retain the lists that vanish on C. The union of their common kernels, computed in the traceless subspace, is the set of D with the required common factor. Similarity preserves the trace moments and cyclicity.

3. Check that each union equals the whole traceless subspace.

There are 7 and 20 representatives in orders three and four, respectively. Every union is full, covering $7 \cdot 256 = 1792$ and $20 \cdot 32768 = 655360$ representative pairs. The complete integer-arithmetic implementation is in `verify_shared_factor.py`. The binary masks in it encode all traceless matrices using the entries except the last diagonal entry, which is determined by the trace equation. The mask `orthogonal(row)` lists the matrices annihilated by one moment form. Bitwise intersection and union implement the common-kernel and coverage checks.

B.2 Traceless factors

Lemma B.3. *Let S be a cyclic matrix of order $n \geq 3$, and let T be nonscalar of that order. Then $\operatorname{tr}(SgTg^{-1})$ takes every value in K as g ranges over the invertible matrices.*

Proof. The following construction uses the trace calculation of Nielsen [16]. For any two cyclic classes of order $r \geq 3$, let α, β be their determinants. To prescribe the trace τ of their product, put $d = \tau - \alpha\beta - (r-1)$, and set $L = I_{r-1} + dE_{21}$, $U = I_{r-1} + E_{12}$. Consider matrices of order r of the forms

$$M = \begin{pmatrix} y & b \\ L & 0 \end{pmatrix}, \qquad N = \begin{pmatrix} 0 & U \\ c & x \end{pmatrix},$$

where y, x are rows and b, c are scalars. The first matrix with top row zero is strictly lower triangular with subdiagonal entries one. Its Krylov vectors $e_1, Me_1, \ldots, M^{r-1}e_1$ are independent and remain so when the top row changes. If M_0 has top row zero, the characteristic polynomial of $M_0 + e_1 v$ for a row v is

$$t^r - v \operatorname{adj}(tI_r - M_0)e_1.$$

The coefficients of the vector polynomial $\operatorname{adj}(tI_r - M_0)e_1$ are the Krylov basis vectors of M_0, since $\operatorname{adj}(tI_r - M_0) = \sum_{i=0}^{r-1} t^{r-1-i} M_0^i$. Thus varying the row gives every monic characteristic polynomial. The same argument for N uses its strictly upper triangular base with superdiagonal entries one and the cyclic vector e_r. Determinants give $b = (-1)^{r-1}\alpha$, $c = (-1)^{r-1}\beta$. Direct multiplication then gives $\operatorname{tr}(MN) = bc + \operatorname{tr}(LU) = \tau$.

Take a largest cyclic summand R of T, of order $r \geq 2$, and write T as a block sum $R \oplus T_0$. If $r \geq 3$, choose the first r basis vectors for S along a cyclic string, so the compression of S onto this block is cyclic. Conjugating R within its block gives every trace contribution on that block. Indeed, the preceding construction gives one pair of representatives for each trace, and simultaneous conjugation puts the first representative in the fixed compression of S while preserving the trace. The complementary contribution is fixed.

If $r = 2$, take a cyclic vector v for S and choose distinct scalars $a, b \in K$. The vectors v, Sv, S^2v are independent since $n \geq 3$. On the plane spanned by $v_1 = (S - bI)v$, $v_2 = (S - aI)v$, choose the complementary space to contain $(S - aI)(S - bI)v$, which is outside the plane by its S^2v coefficient. Both $(S - aI)v_1$ and $(S - bI)v_2$ equal this complementary vector. Hence the compression of S on that plane is $\operatorname{diag}(a, b)$. A nonscalar matrix R of order two has a conjugate with arbitrary first diagonal entry u, by taking a vector z with z, Rz independent and prescribing the first coordinate functional to take values $1, u$ on them. The trace contribution $(a - b)u + b \operatorname{tr} R$ runs over K. This completes the proof. □

Proposition B.4. *Let $n \geq 3$ and $a, b \in K$. The values of $a[A, X][A, Y] + b[A, Y][A, X]$ on $M_n(K)$ fill $M_n(K)$ if $a + b \neq 0$, and $\mathfrak{sl}_n(K)$ if $a = -b \neq 0$.*

Proof. By Lemma B.2, it suffices to realize the targets as $aCD + bDC$ with both C, D traceless. Suppose first $a + b \neq 0$. We find a cyclic, invertible, traceless C_0 for which $D \mapsto aC_0D + bDC_0$ is invertible. If $ab = 0$, take a cyclic permutation matrix of order n. Otherwise let $r = -b/a \neq 1$ and choose the companion of $t^n - 1$ if $r^n \neq 1$, or of $t^n + t + 1$ if $r^n = 1$. Both companions have trace zero and are invertible. For their characteristic polynomial f and any root λ, one has $f(r\lambda) \neq 0$: in the first case it is $r^n - 1$, and in the second it is $(r - 1)\lambda$. Over a splitting field take C_0 upper triangular with diagonal $\lambda_1, \ldots, \lambda_n$. Order the matrix units by increasing row and then decreasing column. In the image of E_{ij}, the left multiplication terms besides $\lambda_i E_{ij}$ have smaller row, and the right multiplication terms besides $\lambda_j E_{ij}$ have larger column. The operator is triangular in this order with diagonal $a\lambda_i + b\lambda_j$. None is zero, so the operator is invertible over K.

For a nonscalar target T, Lemma B.3 applied to the cyclic matrix C_0^{-1} gives a conjugate C of C_0 with $\operatorname{tr}(C^{-1}T) = 0$. Solve $aCD + bDC = T$ using the invertible operator. Taking the trace after multiplication on the left by C^{-1} gives $(a + b) \operatorname{tr} D = 0$, hence $\operatorname{tr} D = 0$. For

$T = \lambda I_n$, instead take a cyclic permutation matrix C and $D = \frac{\lambda}{a+b} C^{-1}$, which are traceless and have the required product.

For $a = -b \neq 0$, scale the target by a^{-1}. A nonscalar traceless target has a commutator expression $[C, D_0]$ with C traceless and regular semisimple by Lemma 2.1. On $K[C]$ the trace pairing is nondegenerate, because it becomes the diagonal trace pairing over a splitting field. Therefore there is a polynomial in C with trace $-\operatorname{tr} D_0$. Adding it to D_0 makes the second factor traceless and preserves the commutator. For a scalar target λI_n of trace zero with $\lambda \neq 0$, use the upper shift $C = J_n$ and $D = \sum_{i=1}^{n-1} i\lambda E_{i+1,i}$. Then $[C, D] = \lambda I_n$ because $n = 0$ in K. Both inputs are traceless, and zero is attained by taking a zero input. Cyclicity of trace gives the reverse inclusion. □

C Finite matrix sources for the standard quartic

For integers a, b with $1 \leq a < b \leq n$ in matrix order n, set $J_{a,b} = \sum_{i=a}^{b-1} E_{i+1,i}$. The matrices N_+, N_- are the two off-diagonal identity blocks from Lemma 12.9. Write F_d for the order-d Fiedler companion in (14). The notation $I_3 \oplus F_d$ with trace parameter one means the constant matrix and the $d-1$ coefficient directions in that Fiedler family.

Each entry below lists its targets and the matrices of its triple, encoded by rows with one matrix in each column, over the indicated prime field. For a single triple P, the square operator is $L_{P,P}$. Entry 10 uses the rectangular operator $L_{P,Q}$ for the triples P, Q in entries 7 and 8, respectively. Each rank in the table is both the rank of the operator and the rank after adjoining the flattened target matrices as extra columns. Thus the targets belong to the image. For entry 1 the listed span includes the whole hollow Fiedler family. For entries 12 and 13 use the member with all free coefficients zero as the first target, and the coefficient matrices of the remaining entries as the other targets.

Here is an entrywise verification of the rank computation. Flatten each matrix row by row. For each matrix unit E_{uv}, expand the column using the 24 signed products in the rectangular prescription of Lemma 12.5. Adjoin the flattened targets to the right of these columns. Read the original columns from left to right. In each column choose the first nonzero entry among rows not yet used, move its row to the next pivot position, divide it by its pivot, and subtract multiples from the rows below. Skip a column if there is no such entry. The numbers of pivots in the original columns are recorded below. In each case the remaining rows are zero in the adjoining columns. All operations here are in the indicated prime field. Since the computations use only the prime field, the image containments hold over every extension of that prime field. For the padding core of entry 11 the cubic value is

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

The displayed matrix has determinant one over $\mathbf{F}_2$. The three maps on 4-by-2 matrices to the triples of Lemma 12.3 have rank eight by the same calculation. Their reverse maps are invertible by trace duality.

Entry 1 $E_{12}, E_{32}, E_{34}, E_{13} + E_{42}, E_{21}$.

```
1000  1001  0001
0101  0101  0100
1001  1110  0000
0001  0000  0000
```

Entry	Characteristic	Rank	Source
1	2	8	Fiedler order four
2	2	30	Cyclic order six
3	2	32	Pattern $(1, 5)$
4	2	32	Pattern $(1, 1, 4)$
5	2	32	Pattern $(2, 4)$
6	2	32	Minimal degree at most three
7	2	12	Scalar and shift in order four
8	2	32	Scalar and shift in order six
9	3	32	Scalar and shift in order six
10	2	23	Order-four to order-six rectangle
11	2	12	Padding core
12	2	44	Padding with cyclic order four
13	2	60	Padding with cyclic order five

Table 4: Ranks of fixed fourth-slot maps and of their augmentations.

Entry 2 $J_{1,6}, E_{12}, E_{13}, E_{14}, E_{15}, E_{16}$.

```
000011  011101  010011
011100  011110  000101
001101  010010  001101
010110  001000  001010
011101  000100  011101
000001  010110  011100
```

Entry 3 $E_{11} + E_{22}, J_{2,6}, E_{23}, E_{24}, E_{25}, E_{26}$.

```
111000  100100  000000
011101  110111  001110
001111  000011  000010
001111  101101  001000
101011  000100  101011
100000  101100  101000
```

Entry 4 $E_{11} + E_{22}, J_{3,6}, E_{34}, E_{35}, E_{36}$.

```
100101  111000  001110
110111  010001  000011
111001  101001  001011
110110  100111  000111
010011  110101  010011
000010  010000  110110
```

Entry 5 $J_{1,2} + J_{3,6}, E_{11} + E_{33}, E_{12}, E_{34}, E_{35}, E_{36}$.

```
011001  100000  001001
000101  000000  010001
000110  001100  010110
110110  010000  100110
110001  110011  110000
010111  100011  000011
```

Entry 6 $\mathrm{diag}(E_{ij}, E_{ij})$ $(1 \le i, j \le 3)$, $E_{12}, E_{21}, E_{11}+E_{22}, E_{23}, N_+, N_-$.

```
101110  001101  001100
011111  000110  001110
100100  100110  100010
100101  101001  010001
101011  101000  100001
111100  111100  000100
```

Entry 7 $I_4, E_{12}, E_{23}, E_{34}, E_{41}$.

```
1010  0110  0000
0100  0011  0010
1001  1000  1010
0100  1100  1101
```

Entry 8 $I_6, E_{i,i+1}$ $(1 \le i < 6), E_{61}$.

```
111010  010010  100110
010001  001010  011011
101110  100100  101001
010100  100100  010100
110011  000011  010011
001101  111100  000001
```

Entry 9 $I_6, E_{i,i+1}$ $(1 \le i < 6)$.

```
220012  200221  122002
012200  012202  022211
002202  200001  200100
120200  200010  000011
211020  211001  020021
021000  002200  201101
```

Entry 10 E_{41} in the 4-by-6 rectangle.

Entry 11 $E_{11}+E_{22}, E_{33}+E_{44}, E_{34}$.

```
1111  0100  1110
1101  0001  0100
1100  1000  0001
1100  0100  0000
```

Entry 12 $I_3 \oplus F_4$ with trace parameter one.

```
0011101  0001011  1001111
0011001  0011101  1001001
0000111  0100000  1010110
1101101  0101110  1001011
1100000  0000001  0100101
0000110  0101001  0000111
0101101  0100001  0111100
```

Entry 13 $I_3 \oplus F_5$ with trace parameter one.

```
00100011  10010000  10010111
10110011  01000111  11101010
11000001  11111100  01000110
01100010  11001110  00011011
01100010  10000010  11101001
01011011  11110011  10100100
01101000  11000000  10101000
11010010  00101010  11110001
```

C.1 Integer determinants

For the triples P_0, Q_1, Q_2, Q_3 of Lemma 12.5, apply its column prescription over $\mathbf{Z}$. The following table gives the first-nonzero pivot in each successive column under Gaussian elimination over $\mathbf{Q}$, with the number of row swaps. For a square operator the determinant is the product of these pivots times the sign of the swaps.

(P, Q)	pivots	swaps
(P_0, Q_1)	$1, 1, 1, 1, 1, 1$	0
(P_0, Q_2)	$1, 2, \frac{1}{2}, 2, 1, \frac{1}{2}$	0
(P_0, Q_3)	$1, 1, 1, 1, 1, 1$	0
(Q_1, Q_2)	$-1, -1, 1, 1$	2
(Q_1, Q_3)	$-1, 1, -1, 1$	2
(Q_2, Q_3)	$-1, 1, -1, 1$	0

The determinants are therefore one as integers and remain one over every K. Multiplying the six cubic words gives $I_3 + E_{13}$ for P_0, and

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \operatorname{diag}(1, -1)$$

for Q_1, Q_2, Q_3, respectively.

For the order-four map of Lemma 12.9 use rows $1, 2, 3, 4, 5, 6, 7, 9, 10, 11$ and columns $1, 2, 3, 5, 6, 7, 9, 10, 11, 13$ in row-major coordinates. Applying the bracket formula to the ten column units gives the minor

$$\begin{pmatrix} 0 & 0 & 0 & -2 & -1 & 0 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 0 & -3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & -2 & -1 & 0 & 0 \\ 1 & 0 & 0 & 3 & 0 & 0 & 0 & -3 & 0 & 0 \\ -1 & 1 & 0 & 0 & 3 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

At elimination step k, let m_{ij} be the current entries after any row swap to place the pivot at (k, k). Replace each trailing entry with $i, j > k$ by $(m_{kk} m_{ij} - m_{ik} m_{kj})$ divided by the pivot used at step $k - 1$, using divisor one at $k = 1$. This fraction-free elimination uses pivots $2, 4, -4, 8, 4, -8, -8, 12, 16$ and gives final entry 64. There are five row swaps, so the integer determinant is -64.

References

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

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

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

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

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

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

[7] M. Fiedler, *A note on companion matrices*, Linear Algebra Appl. **372** (2003), 325–331.

[8] V. T. Filippov, V. K. Kharchenko and I. P. Shestakov (compilers), *Dniester Notebook. Unsolved problems in the theory of rings and modules*, 4th ed., Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 1993. English translation by M. R. Bremner and M. V. Kochetov, in *Non-associative Algebra and Its Applications*, Lecture Notes in Pure and Appl. Math. **246**, Chapman & Hall/CRC, 2006, 461–516.

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

[10] I. Gonzales Gargate and T. C. de Mello, *A new approach to the Lvov–Kaplansky conjecture through gradings*, Linear Algebra Appl. **674** (2023), 453–465. arXiv:2210.05653.

[11] 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.

[12] I. Kaplansky, *Problems in the theory of rings*, in *Report of a Conference on Linear Algebras*, National Academy of Sciences–National Research Council, 1957, 1–3.

[13] S. Malev, *The images of non-commutative polynomials evaluated on* 2×2 *matrices over an arbitrary field*, arXiv:1310.8563v4.

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

[15] Y. T. Miah, *The L'vov–Kaplansky conjecture for standard polynomials*, Hexagon 2611.00184v1 (2026).

[16] K. Nielsen, *On the traces of the product of 2 linear similarity classes*, arXiv:2605.20792 (2026).

[17] S. Pan, *The L'vov–Kaplansky conjecture for multilinear polynomials of degrees four to six*, Hexagon 2611.00173v1 (2026).

[18] K. Shoda, *Einige Sätze über Matrizen*, Japan. J. Math. **13** (1936), 361–365.

[19] A. Stasinski, *Commutators of trace zero matrices over principal ideal rings*, Israel J. Math. **228** (2018), 211–227.

[20] Q. Tang, *The L'vov–Kaplansky conjecture for polynomials of degree four*, Hexagon 2611.00119v1 (2026).

[21] Q. Tang, *The L'vov–Kaplansky conjecture in degrees five and six*, aiXiv:2611.00692v1 (2026).

[22] D. Vitas, *The L'vov–Kaplansky conjecture for polynomials of degree three*, Linear Algebra Appl. **733** (2026), 205–232.

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