Paper accepted by Linear and Multilinear Algebra in 2017

Generalized commutators and a problem related to the Amitsur-Levitzki theorem

John D. Dixon Irwin S. Pressman

June 23, 2017

Abstract

The generalized commutator $[A_1|...|A_k]$ of a list $A_1, ..., A_k$ of k real $n \times n$ matrices is defined as a multilinear skew function and the linear operator $T = T(A_1, ..., A_k)$ on the vector space $M_n(\mathbb{R})$ is defined by $TX := [A_1|...|A_k|X]$. The Amitsur-Levitzki theorem shows that $T = 0$ when $k \geq 2n-1$. We investigate the kernel of T and prove that for all integers k and n such that $2 \leq k \leq 2n-2$ we have $\dim T(A_1, ..., A_k) \geq \nu_0(n,k)$ where $\nu_0(n,k) := k$ if k is even; $k+1$ if k is odd and n is even; and $k+2$ if k and n are both odd. We conjecture that this result is best possible and that $\dim T(A_1, ..., A_k) = \nu_0(n,k)$ for almost all $A_1, ..., A_k$ when k and n are in this range. This conjecture is supported by some computational evidence but , so far remains open.

1 The generalized commutator $[A_1|...|A_k]$

Let $M_n(K)$ be the ring of $n \times n$ matrices over a commutative ring K. In their foundational paper [2] Amitsur and Levitzki considered

$$S_{2n}(A_1, ...A_{2n}) := \sum_{\pi} sgn(\pi) A_{\pi(1)} ... A_{\pi(2n)}$$

where the sum is over all permutations π of $[1, 2, ..., 2n]$ and showed that it is equal to 0 for all $A_1, ..., A_{2n} \in M_n(K)$. For other proofs see [7], [8], [9], [10], [11] and [13].

For all positive integers k and n and $A_1, ..., A_k \in M_n(K)$ we shall use a modification of the notation of [7] and write

$$[A_1|...|A_k] := \sum_{\pi} sgn(\pi) A_{\pi(1)} ... A_{\pi(k)} \quad (1)$$

where the sum is over all permutations π of $[1, 2, ..., k]$. We call $[A_1|...|A_k]$ a *generalized commutator*. It is readily seen that the function $(A_1, ..., A_k) \longmapsto [A_1|...|A_k]$ is multilinear and skew symmetric (compare Lemma 5 below).

Consider the linear operator $T := T(A_1, ..., A_k)$ on $M_n(K)$ defined by $TX := [A_1|...|A_k|X]$. The object of this paper is to investigate properties of T, particularly its kernel $V(A_1, ..., A_k) := \ker T$. When $k = 2n-1$ then $T = 0$ by the Amitsur-Levitzki theorem, and a simple induction argument using (2) below shows that $T = 0$ for all $k \geq 2n-1$, so we can restrict ourselves to the case where $k \leq 2n-2$. On the other hand, if $k = 1$ then $\ker T$ is equal to the centralizer of A_1 and so is a well studied subspace. In the remainder of this paper we shall show that for the other values of k there is evidence that the following is true.

Conjecture 1 *Suppose that $2 \leq k \leq 2n-2$. Then for almost all choices of $A_1, ..., A_k \in M_n(\mathbb{R})$ the dimension d of the kernel $V(A_1, ..., A_k)$ is given by $d = k$ if k is even, $d = k+1$ if k is odd and n is even, and $d = k+2$ if both k and n are odd.*

Remark 2 *We explain what we mean by "almost all" in Section 4. If the conjecture is true then it seems very likely that it holds for arbitrary infinite fields of characteristic $\neq 2$. However, not all the arguments carry through directly from $\mathbb{R}$, so as a first step it is reasonable to attempt to verify the conjecture for the real field.*

2 Properties of the operator $T(A_1, ..., A_k)$

In what follows we shall assume that $K = \mathbb{R}$. If we collect together the products in the sum in (1) which begin with the same factor we obtain

$$[A_1|...|A_k] := \sum_{i=1}^{k} (-1)^{i-1} A_i C_i \tag{2}$$

where $C_i := [A_1|...|\hat{A}_i|...|A_k]$ is the generalized commutator of $k-1$ matrices omitting A_i (we define the empty generalized commutator $[\,] := I$, the identity matrix). For example, we have $[A_1] = A_1$, $[A_1|A_2] = A_1A_2 - A_2A_1$ and

$$[A_1|A_2|A_3] = A_1(A_2A_3 - A_3A_2) - A_2(A_1A_3 - A_3A_1) + A_3(A_1A_2 - A_2A_1).$$

The function $(A_1, ..., A_k) \longmapsto [A_1|...|A_k]$ is linear in each of its arguments. It is skew in the sense that interchanging two of the arguments changes the sign of the generalized commutator and is 0 if two of its arguments are equal. More generally, if arguments are linearly dependent, then some A_j is a linear combination of the other A_i and so $[A_1|...|A_k]$ can be expanded as a linear combination of generalized commutators each of which has two equal arguments. Thus $[A_1|...|A_k] = 0$ whenever the arguments are linearly dependent.

Lemma 3 *Suppose that $A_1, ..., A_k$ and $B_1, ..., B_k$ are two lists of matrices from $M_n(\mathbb{R})$. If there exists a $k \times k$ matrix $C = [\gamma_{ij}]$ such that*

$$B_i := \sum_{j=1}^{k} \gamma_{ij} A_j \text{ for } i = 1, ..., k.$$

Then $[B_1|...|B_k] = (\det C)[A_1|...|A_k]$.

Proof. Since the generalized commutator is multilinear we have

$$[B_1|...|B_k] = \sum \gamma_{1j_1}...\gamma_{kj_k}[A_{j_1}|...|A_{j_k}]$$

where the sum is over all $(j_1, ..., j_k) \in [1, ..., k]^k$. Since a generalized commutator with at least two equal arguments is 0 we can restrict the last sum to k-tuples of the form $(j_1, ..., j_k) = (\pi(1), ..., \pi(k))$ where π runs over all permutations of $[1, ..., k]$. Thus

$$[B_1|...|B_k] = \sum_{\pi} \gamma_{1\pi(1)}...\gamma_{k\pi(k)}[A_{\pi(1)}|...|A_{\pi(k)}] = (\det C)[A_1|...|A_k]$$

Since $[A_{\pi(1)}|...|A_{\pi(k)}] = sgn(\pi)[A_1|...|A_k]$ by the skew property. ■

Remark 4 *The transformation given in the lemma reflects the property that there is a factorization of the linear mapping defined by the generalized commutator through the exterior product. More precisely there are linear mappings* $M_n(\mathbb{R})^k \to \bigwedge^k M_n(\mathbb{R}) \to M_n(\mathbb{R})$ *given by* $(A_1, ..., A_k) \longmapsto A_1 \wedge ... \wedge A_k \longmapsto [A_1|...|A_k]$ *because the generalized commutator is multilinear and skew.*

Write $Sub(A_1, ..., A_k)$ to denote the subspace of $M_n(\mathbb{R})$ spanned by $A_1, ..., A_k$.

Lemma 5 *Given k matrices $A_1, ..., A_k \in M_n(\mathbb{R})$ we have:*
(a) $Sub(A_1, ..., A_k) \subseteq V(A_1, ..., A_k)$;
(b) if $A_1, ..., A_k$ are linearly dependent, then $T(A_1, ..., A_k) = 0$, so $V(A_1, ..., A_k) = M_n(F)$;
(c) if $A_1, ..., A_k$ are linearly independent, then $V(A_1, ..., A_k)$ depends only on the subspace $Sub(A_1, ..., A_k)$ and not on a particular basis;
(d) if k is odd, then $V(A_1, ..., A_k)$ contains the centralizer of $Sub(A_1, ..., A_k)$.

Proof. (a) If $X = A_j$ then $TX = [A_1|...|A_k|X] = 0$ because the generalized commutator has a repeated argument. Thus each $A_j \in V(A_1, ..., A_k)$ and (a) follows.

(b) If $A_1, ..., A_k$ are linearly dependent then $A_1, ..., A_k, X$ are linearly dependent and so $[A_1|...|A_k|X] = 0$ for all X.

(c) If $B_1, ..., B_k$ is a second basis for $Sub(A_1, ..., A_k)$, then by Lemma 3 there is an invertible $(k + 1) \times (k + 1)$ matrix C of the form

$$C = \begin{bmatrix} C_0 & 0 \\ 0 & 1 \end{bmatrix} \text{ where } C_0 \text{ is an invertible } k \times k \text{ block}$$

such that $[B_1|...|B_k|X] = (\det C)[A_1|...|A_k|X]$ for all $X \in M_n(\mathbb{R})$. Hence $V(B_1, ..., B_k) = V(A_1, ..., A_k)$.

(d) Suppose k is odd and that A_{k+1} lies in the centralizer of $Sub(A_1, ..., A_k)$. We have to show that $[A_1|...|A_k|A_{k+1}] = 0$. To do this we classify the permutations π of $[1, ..., k + 1]$ into two classes, Π_1 and Π_2, according to whether the integer $\pi^{-1}(k+1)$ is odd or even. Since $k+1$ is even, $|\Pi_1| = |\Pi_2|$ and we have a bijection $\Pi_1 \to \Pi_2$ defined as follows. If $\pi \in \Pi_1$ then by definition there exists an odd integer i such that $\pi(i) = k+1$. Since $k+1$ is even, $i+1 \leq k+1$ and so we can define a permutation π' by $\pi'(i) = \pi(i+1)$, $\pi'(i+1) = \pi(i)$ $(= k+1)$ and $\pi'(j) = \pi(j)$ for all $j \neq i, i+1$. Clearly $\pi' \in \Pi_2$ and it is readily verified that the mapping $\pi \longmapsto \pi'$ is a bijection of Π_1 onto Π_2 since $|\Pi_1| = |\Pi_2|$. Now $sgn(\pi) = -sgn(\pi')$ and $A_{\pi(i)} = A_{\pi'(i+1)} = A_{k+1}$ centralizes $Sub(A_1, ..., A_k)$ by hypothesis, so we have $sgn(\pi)A_{\pi(1)}...A_{\pi(i)}A_{\pi(i+1)}...A_{\pi(k+1)} + sgn(\pi')A_{\pi'(1)}...A_{\pi'(i)}A_{\pi'(i+1)}...A_{\pi'(k+1)} = 0$. Thus in the expansion of the type (1) for $[A_1|...|A_k|\, A_{k+1}]$ the terms in the sum can be collected in mutually cancelling pairs, and so $[A_1|...|A_k|\, A_{k+1}] = 0$ are required. ■

Remark 6 *1. If k is even and A_{k+1} centralizes $Sub(A_1, ..., A_k)$, then it follows from (2) that $[A_1|...|A_k|A_{k+1}] = \sum_{i=1}^{k+1}(-1)^{i-1}A_iC_i = (-1)^k A_{k+1}[A_1|...|A_k]$ since $C_i = 0$ for each $i \neq k+1$ by part (d) of the lemma.*

2. Since $[A_1|...|A_k|X]' = [X'|A'_k|...|A'_1] = \pm[A'_1|...|A'_k|X']$ where $'$ denotes the transpose, the subspace $V(A'_1, ..., A'_k)$ consists of the transposes of the matrices in $V(A_1, ..., A_k)$.

3 The matrix for $T(A_1, ..., A_k)$

The operator $T = T(A_1, ..., A_k)$ acts on the n^2-dimensional space $M_n(\mathbb{R})$. We describe a matrix for T over the standard basis of $M_n(\mathbb{R})$ in terms of Kronecker products.

For each sublist $\Lambda = i_1 < ... < i_s$ of $[1, 2, ..., k]$ define $c(\Lambda) := [A_{i_1}|...|A_{i_s}]$ and denote the complementary sublist of Λ by $\bar{\Lambda}$. Then we can rewrite the generalized commutator

$$[A_1|...|A_k|X] = \sum_{\Lambda} \sigma(\Lambda)c(\bar{\Lambda})Xc(\Lambda)$$

where the sum is over all 2^k sublists Λ of $[1, 2, ..., k]$ and $\sigma(\Lambda)$ is the sign of the permutation $\lambda : [1, 2, ..., k+1] \longmapsto [\bar{\Lambda}, (k+1), \Lambda]$. For example, if $k = 2$ then

$$[A_1|A_2|X] = [A_1|A_2]X - [A_1]X[A_2] + [A_2]X[A_1] + X[A_1|A_2].$$

Let E_{ij} be the $n \times n$ matrix whose (i,j)th entry is 1 and whose remaining entries are 0. For each $C \in M_n(\mathbb{R})$ we define $vec(C)$ to be the n^2-column vector whose entries represent C in terms of the basis $E_{11}, E_{21}, ..., E_{n1}, ..., E_{1n}, E_{2n}, ..., E_{nn}$ (so $vec(C)$ is obtained by stacking the successive columns $c_1, ..., c_n$ of C). It is known (see, for example, [5, Sect. 4.3]) that $vec(AXB) = (B' \otimes A)vec(X)$ where B' is the transpose of $B = [\beta_{ij}]$ and the Kronecker product $\otimes$ is given by

$$B' \otimes A = \begin{bmatrix} \beta_{11}A & \beta_{21}A & ... & \beta_{n1}A \\ \beta_{12}A & \beta_{22}A & ... & \beta_{n2}A \\ \vdots & \vdots & & \vdots \\ \beta_{1n}A & \beta_{2n}A & ... & \beta_{nn}A \end{bmatrix}.$$

Thus the expression above for $[A_1|...|A_k|X]$ shows that the $n^2 \times n^2$ matrix

$$M = M(A_1, ..., A_k) := \sum_{\Lambda} \sigma(\Lambda) c(\Lambda)' \otimes c(\bar{\Lambda}) \tag{3}$$

satisfies

$$vec([A_1|...|A_k|X]) = Mvec(X)$$

and hence M is the matrix for T over the given basis.

Next note that $\sigma(\Lambda) = \sigma(\bar{\Lambda})$ or $-\sigma(\bar{\Lambda})$ according to whether the permutation which takes $[\bar{\Lambda}, k+1, \Lambda]$ to $[\Lambda, k+1, \bar{\Lambda}]$ is even or odd. If $|\Lambda| = s$ then the permutation which maps $[\bar{\Lambda}, k+1, \Lambda]$ onto $[k+1, \Lambda, \bar{\Lambda}]$ can be obtained by $(k-s)(s+1)$ interchanges, and similarly the permutation which maps $[k+1, \Lambda, \bar{\Lambda}]$ onto $[\Lambda, k+1, \bar{\Lambda}]$ can be obtained with s interchanges. This shows that $\sigma(\Lambda) = \sigma(\bar{\Lambda})$ or $-\sigma(\bar{\Lambda})$ according to whether $(k-s)(s+1)+s$ is even or odd. Hence $\sigma(\Lambda) = \sigma(\bar{\Lambda})$ if k and s are both even; otherwise $\sigma(\Lambda) = -\sigma(\bar{\Lambda})$.

Now [5, Cor. 4.3.10] shows that there is an $n^2 \times n^2$ permutation matrix $P \in M_{n^2}(\mathbb{R})$ such that $P = P^{-1} = P'$ and $P'(A \otimes B)P = B \otimes A$ for every pair (A, B) of $n \times n$ matrices; moreover P is unique and in terms of the basis above is given by $P := \sum_{i=1}^{n} \sum_{j=1}^{n} E_{ij} \otimes E_{ij}^{'}$ (an $n \times n$ block matrix whose (i, j)th block equals E_{ij}'). Thus

$$P'MP = \sum_{\Lambda} \sigma(\Lambda) c(\bar{\Lambda}) \otimes c(\Lambda)' = \left(\sum_{\Lambda} \sigma(\Lambda) c(\bar{\Lambda})' \otimes c(\Lambda) \right)'.$$

Since $M = \sum_{\Lambda} \sigma(\bar{\Lambda}) c(\bar{\Lambda})' \otimes c(\Lambda)$ (replacing Λ by $\bar{\Lambda}$), it follows that

$$P'MP = -M' \text{ if } k \text{ is odd.} \tag{4}$$

4 Generic matrices

A list $A_1, ..., A_k$ of matrices in $M_n(\mathbb{R})$ is called *generic* if the kn^2 entries in these matrices are algebraically independent over $\mathbb{Q}$ (compare [3]). Similarly a list of column vectors $a_1, ..., a_k \in \mathbb{R}^m$ is generic if their km entries are algebraically independent over $\mathbb{Q}$. We observe that if $k \leq m$ then a generic $m \times k$ matrix B has rank k since the determinant of the $k \times k$ submatrix formed from the first k rows of B is a nonzero polynomial in the entries. It follows that when $k \leq m$ every generic list of k vectors in $\mathbb{R}^m$ is linearly independent.

Let $A_1, ..., A_k \in M_n(\mathbb{R})$ be a generic list of matrices and Φ be the set of entries of these matrices. Each mapping $\Phi \to \mathbb{R}$ is called a *specialization*. If $\tilde{A}_1, ..., \tilde{A}_k$ be another list in $M_n(\mathbb{R})$ of the same length, then the specialization defined by $A_i \longmapsto \tilde{A}_i$ $(i = 1, ..., k)$ defines a unique $\mathbb{Q}$-algebra homomorphism of $\mathbb{Q}[A_1, ..., A_k]$ onto $\mathbb{Q}[\tilde{A}_1, ..., \tilde{A}_k]$. This homomorphism is an isomorphism if $\tilde{A}_1, ..., \tilde{A}_k$ is also a generic list since the inverse mapping is also a $\mathbb{Q}$-homomorphism. Let $\nu(n, k)$ be the dimension of the kernel of $T(A_1, ..., A_k)$ (clearly $\nu(n, k)$ is independent of the particular choice of generic matrices).

The matrix $M(A_1, ..., A_k)$ defined in (3) has entries in the polynomial ring $\mathbb{Q}[\Phi]$ and has rank $r := n^2 - \nu(n, k)$. This means that each $(r+1) \times (r+1)$ submatrix of $M(A_1, ..., A_k)$ has determinant 0 but there exists at least one $r \times r$ submatrix with nonzero determinant $\Delta(A_1, ..., A_k) \in \mathbb{Q}[\Phi]$. Since the specialization $A_i \longmapsto \tilde{A}_i$ $(i = 1, ..., k)$ maps $M(A_1, ..., A_k)$ onto $M(\tilde{A}_1, ..., \tilde{A}_k)$ the rank of $M(\tilde{A}_1, ..., \tilde{A}_k)$ is at most r for all $\tilde{A}_1, ..., \tilde{A}_k \in M_n(\mathbb{R})$. Moreover a sufficient condition for its rank to equal r is given by $\Delta(\tilde{A}_1, ..., \tilde{A}_k) \neq 0$. This shows that the dimension of the kernel of $T(\tilde{A}_1, ..., \tilde{A}_k)$ is at least $\nu(n, k)$, and that it is equal to $\nu(n, k)$ whenever $\Delta(\tilde{A}_1, ..., \tilde{A}_k) \neq 0$ holds. Note that $\Delta(A_1, ..., A_k)$ is a $\mathbb{Q}$-polynomial expression in the entries of the A_i. Summing up we have the following facts about the dimension of $V(A_1, ..., A_k) = \ker T(A_1, ..., A_k)$.

Lemma 7 *For all positive integers k and n there exists an integer $\nu(n, k)$ and a nonzero rational polynomial ψ in kn^2 variables such that:*

(a) the dimension of $V(A_1, ..., A_k)$ is at least $\nu(n, k)$ for each list $A_1, ..., A_k$ of length k in $M_n(\mathbb{R})$;

(b) the dimension of $V(A_1, ..., A_k)$ is exactly $v(n, k)$ if $A_1, ..., A_k$ is a generic list of matrices;

(c) the dimension of $V(A_1, ..., A_k)$ is exactly $\nu(n, k)$ whenever the value of ψ is nonzero for the list of entries of $A_1, ..., A_k$.

Corollary 8 *Let ψ have total degree d and choose $\varepsilon > 0$. Then for each finite subset S of $\mathbb{R}$ with $|S| > d/\varepsilon$ and random choices of $A_1, ..., A_k$ with entries in S, the probability that $\dim V(A_1, ..., A_k) = \nu(n, k)$ is at least $1 - \varepsilon$. In particular, in this sense, if R is any nonzero subring of $\mathbb{R}$ (necessarily infinite), then $\dim V(A_1, ..., A_k) = \nu(n, k)$ for "almost all" $A_1, ..., A_k \in M_n(R)$.*

Proof. Schwartz [12] shows that, if $\varphi(x_1, ..., x_m)$ is a nonzero polynomial of total degree d over any field F and S is a finite subset of F, then the proportion of points in S^m at which φ vanishes is not greater than $d/|S|$ (similar ideas appear in [14]). Now suppose ψ has total degree d. Then Schwartz's lemma shows that for each $\varepsilon > 0$ and each finite $S \subseteq \mathbb{R}$ with $|S| > d/\varepsilon$, the probability that ψ has a nonzero value at a random point in S^{kn^2} is $> 1 - \varepsilon$. Thus (c) shows that in this sense $\dim V(A_1, ..., A_k) = \nu(n, k)$ for almost all lists $A_1, ..., A_n \in M_n(R)$. The set of exceptions also has Lebesgue measure 0 in $M_n(\mathbb{R})^k$. ∎

We can prove some lower bounds for $\nu(n, k)$.

Lemma 9 *Let n and k be positive integers. Then*

(a) $\nu(n, 1) = n$ for all n;

(b) $\nu(n, k) \geq k$ for $2 \leq k \leq 2n - 2$;

(c) $\nu(n, k) \geq \nu(n, k+1) \geq k + 1$ if k is odd and $3 \leq k \leq 2n - 2$;

(d) $\nu(n, k) \geq k + 2$ if k and n are both odd and $3 \leq k \leq 2n - 2$;

(e) $\nu(n, k) = n^2$ if $k \geq 2n - 1$.

Proof. (a) If $k = 1$ then $T(A_1)X = A_1X - XA_1$ and so $V(A_1)$ is the centralizer of A_1. It is well known that the dimension of the centralizer of an $n \times n$ matrix

$A_1 \in M_n(\mathbb{R})$ is always at least n and it is exactly n if and only if A_1 is a cyclic (= nonderogatory) matrix (see, for example, [4, Sect. 3.2.4]). Hence $\nu(n,1) = n$.

(b) Suppose that $2 \leq k \leq 2n-2$ (so $n \geq 2$). As noted above, if $k \leq m$, then a generic list of k vectors in $\mathbb{R}^m$ is linearly independent; in particular a generic list of k matrices in $M_n(\mathbb{R})$ is linearly independent if $k \leq 2n-2 \leq n^2$. Hence Lemma 5(a) shows that $\nu(n,k) \geq k$.

(c) Suppose that k is odd and $3 \leq k \leq 2n-2$. Then from the remark following Lemma 5 we see that

$$[A_1|...|A_k|I|X] = -[A_1|...|A_k|X|I] = -[A_1|...|A_k|X]$$

because I centralizes $Sub(A_1, ..., A_k, X)$. Taking a generic list $A_1, ..., A_k$ of matrices in $M_n(\mathbb{R})$, we have

$$\nu(n,k) = \dim V(A_1, ..., A_k) = \dim V(A_1, ..., A_k, I) \geq \nu(n, k+1).$$

Thus $\nu(n,k) \geq k+1$ by (b).

(d) Suppose that both n and k are odd with $3 \leq k \leq 2n-2$. Let $A_1, ..., A_k \in M_n(\mathbb{R})$ be a generic list of k matrices and consider $M = M(A_1, ..., A_k)$. Then (4) shows that $MP = -PM' = -(MP)'$ since $P = P^{-1} = P'$. Since P is invertible, the dimension of the nullspace of MP is equal to the dimension $\nu(n,k)$ of the null space of M. The skew symmetric matrix MP is diagonalizable over $\mathbb{C}$ and 0 is its only real eigenvalue. Thus MP has an even number of nonzero eigenvalues. Because MP is diagonalizable, $\nu(n,k)$ is equal to the multiplicity of 0 as an eigenvalue of MP, and therefore $\nu(n,k) \equiv n^2 \pmod 2$. By hypothesis n and $k \geq 3$ are both odd, so $\nu(n,k)$ is odd and $\nu(n,k) \geq k+1$ by (c). Since $k+1$ is even we conclude that $\nu(n,k) \geq k+2$.

(e) This follows from the Amitsur-Levitzki theorem. ∎

Definition 10 *Define* $\nu_0(n,1) := n$ *and* $\nu_0(n,k) := n^2$ *if* $k \geq 2n-1$. *For* $2 \leq k \leq 2n-2$ *define*

$$\nu_0(n,k) := \begin{cases} k & \text{if } k \text{ is even} \\ k+1 & \text{if } k \text{ is odd and } n \text{ is even} \\ k+2 & \text{if } k \text{ is odd and } n \text{ is odd} \end{cases}.$$

Then Lemma 9 shows that $\nu(n,k) \geq \nu_0(n,k)$ for all positive integers n and k, and equality holds for $k=1$ and for $k \geq 2n-1$. Our conjecture (see the Introduction) is that equality holds for all n and k.

Since $\nu_0(n,k)$ is a lower bound for $\dim V(A_1, ..., A_k)$ for all $A_1, ..., A_k \in M_n(\mathbb{R})$, in order to prove the conjecture for a particular pair (n,k) it is enough to show that there is at least one list of length k in $M_n(\mathbb{R})$ such that $\dim V(A_1, ..., A_k) = \nu_0(n,k)$; this shows that $\nu_0(n,k)$ is the greatest lower bound for $V(A_1, ..., A_k)$ and hence equal to $\nu(n,k)$. On the other hand, if the conjecture is true then this equality will hold for "almost all" lists of length k in $M_n(\mathbb{Z})$, for example, so it should not be hard to find suitable $A_1, ..., A_k$. The difficulty lies in proving that $\dim V(A_1, ..., A_k) = \nu_0(n,k)$ for a suitable choice of $A_1, ..., A_k$.

5 Verification of the conjecture for small values of n and k

We wrote simple programs to compute generalized commutators and used these to compute the $n^2 \times n^2$ matrix $M(A_1, ..., A_k)$ given by (3). The complexity of this calculation is dominated by the matrix multiplications and there are approximately $(k+1)!$ of these. The time to multiply two $n \times n$ matrices together using ordinary matrix multiplication is proportional to n^3, so the complexity of computing $M(A_1, ..., A_k)$ is roughly proportional to $n^{3(k+1)!}$. The nullspace for $M(A_1, ..., A_k)$ was computed using a program with complexity roughly proportional to $(n^2)^3 = n^6$. The calculations have to be done using exact arithmetic since the nullspace computation quickly degrades if floating point is used. The computations were carried out independently in MATLAB and J [6].

For $n \leq 8$ and $2 \leq k \leq \min(2n-2, 8)$, we chose random values from the set $\{0, 1, 2\}$ as entries for the matrices A_i and quickly found examples for which $\dim V(A_1, ..., A_k) = \nu_0(n, k)$ (in almost all cases at the first attempt). This illustrates the fact that the estimate in Corollary 8 for the size of S is sometimes excessive. It is not clear how feasible it is to extend these computations since the calculations become much slower as k and n grow. Our results are given in the table below and show that the conjecture in true in this range.

Values of $\nu(n, k)$

n\k	1	2	3	4	5	6	7	8
2	2^c	2	4^{al}	4^{al}	4^{al}	4^{al}	4^{al}	4^{al}
3	3^c	2	5	4	9^{al}	9^{al}	9^{al}	9^{al}
4	4^c	2	4	4	6	6	16^{al}	16^{al}
5	5^c	2	5	4	7	6	9	8
6	6^c	2	4	4	6	6	8	8
7	7^c	2	5	4	7	6	9	8
8	8^c	2	4	4	6	6	8	8

c cyclic matrix al Amitsur-Levitzki Theorem

Remark 11 *As we saw in Lemma 9, if n is even and k is odd and the kernel $V(A_1, ..., A_k)$ of $T(A_1, ..., A_k)$ has dimension $\nu_0(n, k) = k+2$, then $V(A_1, ..., A_k)$ contains $k+1$ linearly independent elements, namely, $A_1, ..., A_k, I$. In general we do not know how to construct a further matrix, say C, to complete a basis for $V(A_1, ..., A_k)$ since the proof in the lemma is only an existence proof. In the calculations we have made for a basis of $V(A_1, ..., A_k)$ in this situation, the matrix C we obtain has no obvious relation to the input (informally we refer to C as a* monster *matrix). It would be interesting to be able to describe the form of this monster matrix.*

6 Bibliography

References

[1] N. Alon, Combinatorial Nullstellensatz, Combin. Probab. Comput. 8 (1999) 7-29.

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

[3] E. Formanek, The ring of generic matrices, Special issue in celebration of Claudio Procesi's 60th birthday, J. Algebra 258 (2002) 310–320.

[4] R.A. Horn and C.R. Johnson, "Matrix Analysis", Cambridge Univ. Press, Cambridge, 1985.

[5] R.A. Horn and C.R. Johnson, "Topics in Matrix Analysis", Cambridge Univ. Press, Cambridge, 1991.

[6] http://www.jsoftware.com/

[7] B. Kostant, A theorem of Frobenius, a theorem of Amitsur-Levitski and cohomology theory, J. Math. Mech. 7 (1958) 237–264.

[8] F. W. Owens, Applications of graph theory to matrix theory, Proc. Amer. Math. Soc. 51 (1975) 242–249.

[9] C. Procesi, On the theorem of Amitzur-Levitzki, Israel J. Math. 207 (2015) 151–154.

[10] Ju. P. Razmyslov, Identities with trace in full matrix algebras over a field of characteristic zero, Izv. Akad. Nauk SSSR. Ser. Mat. 38 (1974), 723–756..

[11] S. Rosset, A new proof of the Amitsur-Levitski identity, Israel J. Math. 23 (1976) 187–188.

[12] J.T. Schwartz, Fast probabilistic algorithms for verification of polynomial identities, J. Assoc. Comput. Mach. 27 (1980) 701–717.

[13] Richard G. Swan, An application of graph theory to algebra, Proc. Amer. Math. Soc. 14 (1963) 367–373; *errata* Proc. Amer. Math. Soc. 21 (1969) 379–380.

[14] R. E. Zippel, Probabilistic Algorithms for Sparse Polynomials, Ph.D. thesis, Massachusetts Institute of Technology, (1979).