Mathematica Pannonica
13/2 (2002), 175–181

FINITE GENERATING SYSTEM OF MATRIX INVARIANTS

M. **Domokos**

Rényi Institute of Mathematics, Hungarian Academy of Sciences, P.O. Box 127, 1364 Budapest, Hungary; Temporary address (until February 2004): School of Mathematics, University of Edinburgh, James Clerk Maxwell Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JZ, Scotland

Dedicated to Richárd Wiegandt on his 70th birthday

Received: May 2002

MSC 2000: 13 A 50, 15 A 72, 16 R 10, 16 R 30

Keywords: Characteristic polynomial, matrix invariant, degree bound, T-ideal, nil algebra.

Abstract: A finite system of generators is determined for the algebra of simultaneous conjugation invariants of m-tuples of $n \times n$ matrices, over an infinite base field of positive characteristic.

Throughout this paper K is an infinite field of characteristic $p > 0$, and let n, m be fixed positive integers with $n \geq 2$. Take an mn^2-variable commutative polynomial algebra

$$K_{n,m} := K[x_{r,ij} \mid 1 \leq i, j \leq n, \; r = 1, \dots, m],$$

and form the generic matrices

$$X_r := (x_{r,ij})_{i,j=1}^{n}, \quad r = 1, \dots, m,$$

contained in the $n \times n$ matrix algebra $M(n, K_{n,m})$ over $K_{n,m}$. The

E-mail address: domokos@renyi.hu
This research was supported through a European Community Marie Curie Fellowship. Partially supported by OTKA No. F 32325 and T 34530.

coefficients $\sigma_j(A)$ of the characteristic polynomial of an $n \times n$ matrix A over a commutative ring are defined by the equality

$$\det(\lambda I - A) = \lambda^n - \sigma_1(A)\lambda^{n-1} + \ldots + (-1)^n \sigma_n(A) I$$

where λ is a commuting indeterminate and I is the identity matrix. Note that σ_1 is the trace and σ_n is the determinant. We may form a product $X_{i_1} \ldots X_{i_s}$ in $M(n, K_{n,m})$, and the characteristic coefficients $\sigma_j(X_{i_1} \ldots X_{i_s})$ $(j = 1, \ldots, n)$ are elements of $K_{n,m}$. Our main object of study is $R_{n,m}$, the unitary K-subalgebra of $K_{n,m}$ generated by

$$(1) \qquad \{\sigma_j(X_{i_1} \ldots X_{i_s}) \mid 1 \leq j \leq n,\ s \in \mathbb{N},\ 1 \leq i_1, \ldots, i_s \leq m\}.$$

Denote by $T_{n,m}$ the unitary $R_{n,m}$-subalgebra of $M(n, K_{n,m})$ generated by $X_1, \ldots, X_m$.

Identify $K_{n,m}$ with the coordinate ring of the space $M_{n,m} := := M(n,K) \oplus \ldots \oplus M(n,K)$ of m-tuples of $n \times n$ matrices, by identifying $x_{r,ij}$ with the function which maps $(A_1, \ldots, A_m)$ to the (i,j)-entry of A_r. Then X_r is identified with the map $M_{n,m} \to M(n,K)$, given by $(A_1, \ldots, A_m) \mapsto A_r$. The general linear group acts on $M(n,K)$ by conjugation and on $M_{n,m}$ by simultaneous conjugation. Donkin proved in [5], [6] that $R_{n,m}$ coincides with the algebra of $GL(n,K)$-invariant polynomial functions $M_{n,m} \to K$. Consequently, from properties of the trace function it follows that $T_{n,m}$ coincides with the $GL(n,K)$-equivariant polynomial maps $M_{n,m} \to M(n,K)$. Therefore $R_{n,m}$ is called the ring of matrix invariants, and $T_{n,m}$ is called the ring of matrix concomitants.

Being the ring of invariants of a reductive group, the algebra $R_{n,m}$ is finitely generated. The aim of the present paper is to determine a finite generating system of $R_{n,m}$. In the case when $\mathrm{char}(K) = 0$ the solution of this problem is well known, we refer to [8] for a survey. An adaptation of a well known argument to the present situation allows us to give an upper bound for s in the generating system (1), see Lemma 1. The resulting finite generating system is refined further with the aid of a new relation between the characteristic coefficients, see Lemma 2.

Let us recall the notion of T-ideals. Take a set $Y_m = \{y_1, \ldots, y_m\}$ of non-commuting variables, and denote by $K\langle Y_m \rangle$ the free associative K-algebra without unity generated by Y_m. The elements of $K\langle Y_m \rangle$ are non-commutative polynomials in the variables $y_1, \ldots, y_m$ with coefficients from K, and having zero constant term. An ideal I of $K\langle Y_m \rangle$ (I is assumed to be a K-subspace) is called a T-ideal, if $f(y_1, \ldots, y_m) \in$

$\in I$, $u_1, \dots, u_m \in K\langle Y_m \rangle$ imply that $f(u_1, \dots, u_m) \in I$. In other words, a T-ideal is an ideal I of $K\langle Y_m \rangle$ which is stable with respect to all K-algebra endomorphisms of $K\langle Y_m \rangle$.

A theorem of Kaplansky [9] asserts that if B is a finitely generated K-algebra such that $x^n = 0$ holds for all $x \in B$, then B is nilpotent. In other words, there exists a natural number N (depending on n, m, K) such that the T-ideal $\{y_1^n\}_T$ of $K\langle Y_m \rangle$ generated by y_1^n contains all words $y_{i_1} \dots y_{i_N}$. Denote by $N(n,m,K)$ the minimal such N, that is,

$$N(n,m,K) :=$$
$$:= \min\{N \mid \forall i_1, \dots, i_N \in \{1, \dots, m\} : \; y_{i_1} \dots y_{i_N} \in \{y_1^n\}_T\}.$$

Elementary linear algebra arguments show that $N(n,m,K) = N(n,m,L)$, if L is another infinite field of characteristic p. Therefore we shall write $N(n,m,p)$ instead of $N(n,m,K)$. An explicit upper bound for $N(n,m,p)$ was given by A. J. Belov [2], who proved that $N(n,m,p) \leq n^6 m^{n+1}$. This bound was improved in [10], showing $N(n,m,p) < \frac{1}{6} n^6 m^n$.

We need to recall the formula of Amitsur [1] expressing the characteristic coefficients of a sum of matrices as a polynomial of characteristic coefficients of products of the summands. Let S be the free semigroup generated by k variables $z_1, \dots, z_k$. We say that $m_1, m_2 \in S$ are equivalent if one is obtained by cyclic permutation from the other. For a word $w = z_{i_1} \dots z_{i_r}$ define the length of w as $l(w) := r$, and $\nu(w) := (\nu_1, \dots, \nu_k) \in \mathbb{N}_0^k$, where ν_j is the number of appearances of z_j in w. We have $l(w) = \nu_1 + \dots + \nu_k$. We call a word indecomposable, if it is not a power of a word of smaller length. Let S_0 be a set of representatives of equivalence classes of indecomposable words. Assume now that $z_1, \dots, z_k$ are $n \times n$ matrices over some commutative ring C, and let $\lambda_1, \dots, \lambda_k$ be commuting indeterminates over C. For $\nu = (\nu_1, \dots, \nu_k) \in \mathbb{N}_0^k$ write $\lambda^\nu := \lambda_1^{\nu_1} \dots \lambda_k^{\nu_k}$. By [1, Th. A] for $j = 1, \dots, n$ we have

$$\sigma_j(\lambda_1 z_1 + \dots + \lambda_k z_k) = \tag{2}$$
$$= \sum (-1)^{j-(i_1+\dots+i_r)} \lambda^{i_1\nu(w_1)+\dots+i_r\nu(w_r)} \sigma_{i_1}(w_1) \dots \sigma_{i_r}(w_r)$$

where the sum ranges over all subsets $\{w_1, \dots, w_r\} \subseteq S_0$ and numbers $i_1, \dots, i_r \in \{1, \dots, j\}$ with $i_1 l(w_1) + \dots + i_r l(w_r) = j$.

The polynomial algebra $K_{n,m}$ is graded in the usual way. The algebra $R_{n,m}$ is a graded subalgebra, since it is generated by homogeneous elements. Denote by R^+ the maximal ideal of $R_{n,m}$ generated by the homogeneous elements of positive degree. A set of homogeneous elements generates $R_{n,m}$ as a K-algebra if and only if its image modulo $(R^+)^2$ generates $R^+/(R^+)^2$ as a K-vector space. We shall call the elements of $(R^+)^2$ decomposable, since they can be removed from any system of homogeneous generators of $R_{n,m}$.

Lemma 1. *The invariant* $\sigma_j(X_{i_1} \dots X_{i_s})$ *is decomposable if* $s > N(n,m,p)$.

Proof. The idea of the proof is well known: it goes back to [7], [11], [12] (see [8] for history). Set $N := N(n,m,p)$, and take an arbitrary monomial $X_{i_1} \dots X_{i_N} Z$ of $X_1, \dots, X_m$ whose degree is greater than N. We shall show that $\sigma_j(X_{i_1} \dots X_{i_N} Z)$ is contained in $(R^+)^2$.

By the definition of N there exists a finite set

$$\{w^\alpha = (w_0^\alpha, w_1^\alpha, w_2^\alpha) \mid \alpha \in \mathcal{A}\}$$

of triples of elements of $K\langle Y_m \rangle$ and coefficients $\{c_\alpha \in K \mid \alpha \in \mathcal{A}\}$ such that

$$y_{i_1} \dots y_{i_N} = \sum_{\alpha \in \mathcal{A}} c_\alpha w_0^\alpha (w_1^\alpha)^n w_2^\alpha \tag{3}$$

holds in $K\langle Y_m \rangle$. Make the substitution $y_i \mapsto X_i$ in (3) to get the equality

$$\begin{aligned} &X_{i_1} \dots X_{i_N} = \\ &= \sum_{\alpha \in \mathcal{A}} c_\alpha w_0^\alpha (X_1, \dots, X_m) w_1^\alpha (X_1, \dots, X_m)^n w_2^\alpha (X_1, \dots, X_m) \end{aligned} \tag{4}$$

in $T_{n,m}$. Apply the Cayley-Hamilton Theorem for the matrix $W_\alpha \in T_{n,m}$, where $W_\alpha := w_1^\alpha(X_1, \dots, X_m)$. Since any characteristic coefficient of W_α is contained in R^+, we conclude that the nth power of W_α is contained in the ideal $R^+ \cdot T_{n,m}$ of $T_{n,m}$. Therefore (4) implies that $X_{i_1} \dots X_{i_N} \in R^+ \cdot T_{n,m}$, hence it can be written as $\sum_i r_i u_i$, where $r_i \in R^+$ and $u_i \in T_{n,m}$. By formula (2) the jth characteristic coefficient of $X_{i_1} \dots X_{i_N} Z = \sum r_i u_i Z$ can be expressed as a polynomial in the elements r_i and the characteristic coefficients of certain products of the elements $u_i Z$, such that in each term at least one r_i and one $u_i Z$ appears. It follows that $\sigma_j(X_{i_1} \dots X_{i_N} Z) \in (R^+)^2$ for all non-trivial monomials Z. ◊

For $d \in \{1, \dots, n\}$ denote by $R_{n,m}(d)$ the unitary K-subalgebra of $R_{n,m}$ generated by all the elements of the form $\sigma_j(X_{i_1} \dots X_{i_s})$ with

$j \leq d$. Denote by $\lfloor \frac{n}{2} \rfloor$ the lower integral part of $\frac{n}{2}$. The following lemma may be viewed as a replacement of the multiplicative property of $\sigma_n = \det$ for certain other characteristic coefficients.

Lemma 2. *As an $R_{n,m}(\lfloor \frac{n}{2} \rfloor)$-algebra, $R_{n,m}$ is generated by the elements*

$$\sigma_j(X_i) \quad \text{with} \quad \lfloor n/2 \rfloor + 1 \leq j \leq n, \ 1 \leq i \leq m.$$

Proof. Observe that the expression (2) is independent from n. So the formula (2) remains valid if $j > n$, and we evaluate $\sigma_{n+1}, \sigma_{n+2}, \ldots$ identically zero. Indeed, for $j > n$ we may identify $M(n, C)$ with the subalgebra of $M(j, C)$ consisting of matrices with zero entries outside the upper left $n \times n$ block. If σ_i $(i = 1, \ldots, j)$ denotes the ith characteristic coefficient function on $j \times j$ matrices, then the restrictions $\sigma_{n+1}\big|_{M(n,C)}, \sigma_{n+2}\big|_{M(n,C)}, \ldots, \sigma_j\big|_{M(n,C)}$ are identically zero, and $\sigma_1\big|_{M(n,C)}, \ldots, \sigma_n\big|_{M(n,C)}$ are the characteristic coefficient functions on $n \times n$ matrices.

Now choose $s \in \{\lfloor \frac{n}{2} \rfloor + 1, \ldots, n\}$. Then $2s > n$, and the special case $j = 2s$ and $k = 2$ of (2) gives the identity

$$(5) \qquad 0 = \sum (-1)^{i_1 + \ldots + i_r} \lambda^{i_1 \nu(w_1) + \ldots + i_r \nu(w_r)} \sigma_{i_1}(w_1) \ldots \sigma_{i_r}(w_r)$$

where the sum ranges over all subsets $\{w_1, \ldots, w_r\} \subseteq S_0$ and natural numbers $i_1, \ldots, i_r \leq n$ with $i_1 l(w_1) + \ldots + i_r l(w_r) = 2s$. It follows that the coefficient of $\lambda_1^s \lambda_2^s$ in (5) is zero, that is, for all $z_1, \ldots, z_k \in M(n, C)$ we have

$$(6) \qquad \begin{aligned} &(-1)^{s+1} \sigma_s(z_1 z_2) - \sigma_s(z_1) \sigma_s(z_2) = \\ &= \sum (-1)^{i_1 + \ldots + i_r} \sigma_{i_1}(w_1) \ldots \sigma_{i_r}(w_r) \end{aligned}$$

where the sum ranges over all subsets $\{w_1, \ldots, w_r\} \subseteq S_0$ and natural numbers $i_1 \ldots, i_r < s$ with $i_1 l(w_1) + \ldots + i_r l(w_r) = 2s$.

Formula (6) implies that if $U, V \in T_{n,m}$ are non-trivial monomials in $X_1, \ldots, X_m$, and $s > \lfloor \frac{n}{2} \rfloor$, then $\sigma_s(UV)$ is contained in the $R_{n,m}(s - 1)$-subalgebra generated by $\sigma_s(U)$ and $\sigma_s(V)$. Applying induction on the degree of U we get that for any monomial U, $\sigma_s(U)$ is contained in the $R_{n,m}(s-1)$-subalgebra generated by $\sigma_s(X_1), \ldots, \sigma_s(X_m)$. Consequently, $R_{n,m}(s)$ is generated as an $R_{n,m}(s-1)$-algebra by $\sigma_s(X_1), \ldots, \sigma_s(X_m)$ for all $s = \lfloor \frac{n}{2} \rfloor + 1, \ldots, n$. This proves Lemma 2. ◊

Combining Lemma 1 and 2 we obtain the following:

Theorem 3. *The algebra $R_{n,m}$ is generated by $\sigma_j(X_i)$, $\sigma_k(X_{i_1} \ldots X_{i_s})$, where $\lfloor \frac{n}{2} \rfloor + 1 \leq j \leq n$, $k \leq \lfloor \frac{n}{2} \rfloor$, $s \leq N(n, m, p)$, $1 \leq i, i_1, \ldots, i_s \leq$*

$\leq m$. *In particular,* $R_{n,m}$ *is generated by its elements of degree* $\leq$ $\leq \lfloor \frac{n}{2} \rfloor N(n,m,p)$.

Proof. Lemma 2 asserts that $R_{n,m}$ is generated by the elements $\sigma_j(X_i)$, $\sigma_k(X_{i_1} \dots X_{i_s})$, where $j \geq \lfloor \frac{n}{2} \rfloor + 1$, $k \leq \lfloor \frac{n}{2} \rfloor$, $s \in \mathbb{N}$, $1 \leq i, i_1, \dots, i_s \leq$ $\leq m$. By Lemma 1 the elements $\sigma_k(X_{i_1} \dots X_{i_s})$ with $s > N(n,m,p)$ are decomposable, hence they can be omitted from this generating system. ◊

Theorem 4. *Assume that char*$(K) = p > \lfloor \frac{n}{2} \rfloor$. *Then* $R_{n,m}$ *is generated by* $\sigma_j(X_i)$, $tr(X_{i_1} \dots X_{i_s})$, *where* $\lfloor \frac{n}{2} \rfloor + 1 \leq j \leq n$, $s \leq N(n,m,p)$, $1 \leq i, i_1, \dots, i_s \leq m$, *and we use the usual notation* $tr := \sigma_1$. *In particular,* $R_{n,m}$ *is generated by its elements of degree* $\leq N(n,m,p)$.

Proof. Since $p > \lfloor \frac{n}{2} \rfloor$, we may apply the Newton formulae in order to express the characteristic coefficients $\sigma_1(A), \dots, \sigma_{\lfloor \frac{n}{2} \rfloor}(A)$ by traces of powers of A for $A \in T_{n,m}$. Hence $R_{n,m}(\lfloor \frac{n}{2} \rfloor)$ is contained in $R_{n,m}(1)$, and by Lemma 2 we get that $R_{n,m}$ is generated by the elements $\sigma_j(X_i)$, $tr(X_{i_1} \dots X_{i_s})$, where $j \geq \lfloor \frac{n}{2} \rfloor + 1$, $s \in \mathbb{N}$, $1 \leq i, i_1, \dots, i_s \leq m$. By Lemma 1 the elements $tr(X_{i_1} \dots X_{i_s})$ with $s > N(n,m,p)$ are decomposable, hence they can be omitted from this generating system. ◊

It is shown in [4] that if $p \leq n$, then $tr(X_1 \dots X_m)$ is not decomposable in $R_{n,m}$. More generally, it is proved in [3] that if $p^r \leq n$ with some positive integer r, then $\sigma_j(X_1 \dots X_m)$ is not decomposable in $R_{n,m}$ for $j = 1, \dots, p^r - 1$. In particular, if $p^r \leq n$ with $r \in \mathbb{N}$, then $R_{n,m}$ is not generated by its elements of degree $< m(p^r - 1)$. The main result of [13] could be applied as well in order to get a lower degree bound for the generators of $R_{n,m}$ in terms of the numbers $N(n,k,p)$.

References

[1] AMITSUR, S. A.: On the characteristic polynomial of a sum of matrices, *Lin. Multilin. Alg.* **8** (1980), 177–182.

[2] BELOV, A. J.: Some estimations for nilpotence of nil-algebras over a field of an arbitrary characteristic and height theorem, *Comm. Alg.* **20**/10 (1992), 2919–2922.

[3] DOMOKOS, M.: Matrix invariants and the failure of Weyl's Theorem, Preprint (2002).

[4] DOMOKOS, M., KUZMIN, S. G. and ZUBKOV, A. N.: Rings of matrix invariants in positive characteristic, *J. Pure Appl. Alg.*, to appear.

[5] DONKIN, S.: Invariants of several matrices, *Inv. Math.* **110** (1992), 389–401.

[6] DONKIN, S.: Invariant functions on matrices, *Math. Proc. Camb. Phil. Soc.* **113** (1993), 23–43.

[7] DUBNOV, J. and IVANOV, V.: Sur l'abaissement de degré des polynômes en affineurs, *C. R. (Doklady) Acad. Sci. URSS* **41** (1943), 95–98.

[8] FORMANEK, E.: The Nagata-Higman Theorem, *Acta Applicandae Mathematicae* **21** (1990), 185–192.

[9] KAPLANSKY, I.: On a problem of Kurosch and Jacobson, *Bull. Amer. Math. Soc.* **52** (1946), 496–500.

[10] KLEIN, A. A.: Bounds for indices of nilpotency and nility, *Arch. Math. (Basel)* **74** (2000), 6–10.

[11] PROCESI, C.: The invariant theory of $n \times n$ matrices, *Adv. Math.* **19** (1976), 306–381.

[12] RAZMYSLOV, Y. P.: Trace identities of full matrix algebras over a field of characteristic 0 (Russian), *Izv. Akad. Nauk SSSR Ser. Mat.* **38** (1974), 723–756.

[13] ZUBKOV, A. N.: On a generalization of the Procesi-Razmyslov Theorem (Russian), *Algebra i Logika* **35** (1996), 433–457.