Serdica Math. J. **32** (2006), 269–276

Serdica
Mathematical Journal

Bulgarian Academy of Sciences
Institute of Mathematics and Informatics

CAYLEY-HAMILTON THEOREM FOR MATRICES OVER AN ARBITRARY RING

Jenő Szigeti*

Communicated by V. Drensky

Abstract. For an $n \times n$ matrix A over an arbitrary unitary ring R, we obtain the following Cayley-Hamilton identity with right matrix coefficients:

$$(\lambda_0 I + C_0) + A(\lambda_1 I + C_1) + \cdots + A^{n-1}(\lambda_{n-1} I + C_{n-1}) + A^n(n!I + C_n) = 0,$$

where $\lambda_0 + \lambda_1 x + \cdots + \lambda_{n-1} x^{n-1} + n! x^n$ is the right characteristic polynomial of A in $R[x]$, $I \in M_n(R)$ is the identity matrix and the entries of the $n \times n$ matrices C_i, $0 \le i \le n$ are in $[R, R]$. If R is commutative, then

$$C_0 = C_1 = \cdots = C_{n-1} = C_n = 0$$

and our identity gives the $n!$ times scalar multiple of the classical Cayley-Hamilton identity for A.

2000 *Mathematics Subject Classification:* 15A15, 15A24, 15A33, 16S50.
Key words: The preadjoint and the right characteristic polynomial of an $n \times n$ matrix, the commutator subgroup $[R, R]$ of a ring R.
*The author was supported by OTKA of Hungary, grant no. T43034 and K61007.

1. Introduction. The Cayley-Hamilton theorem and the corresponding trace identity play a fundamental role in proving classical results about the polynomial and trace identities of the $n \times n$ matrix algebra $M_n(K)$ over a field K. In case of $\mathrm{char}(K) = 0$, Kemer's pioneering work (see [2], [3]) on the T-ideals of associative algebras revealed the importance of the identities satisfied by the $n \times n$ matrices over the Grassmann (exterior) algebra

$$E = K \langle v_1, v_2, \ldots, v_i, \cdots \mid v_i v_j + v_j v_i = 0 \text{ for all } 1 \leq i < j \rangle$$

generated by the infinite sequence of the anticommutative indeterminates $(v_i)_{i \geq 1}$. Accordingly, the importance of matrices over non-commutative rings is an evidence in the theory of PI-rings, nevertheless this fact has been obvious for a long time in other branches of algebra (structure theory of semisimple rings, K-theory, quantum matrices, etc.). Thus a Cayley-Hamilton type identity for such matrices seems to be of general interest.

For $n \times n$ matrices over a Lie-nilpotent ring R a Cayley-Hamilton type identity with one sided scalar coefficients (left or right) was found in [9] (see also in [10]), if R satisfies the PI

$$[[[\ldots[[x_1, x_2], x_3], \ldots], x_m], x_{m+1}] = 0,$$

then the degree of this identity is n^m. Since E is Lie nilpotent with $m = 2$, the above mentioned identity for a matrix $A \in M_n(E)$ is of degree n^2. In [1] Domokos presented a slightly modified version of this identity in which the coefficients are invariant under the conjugate action of $GL_n(K)$.

In the general case (when R is an arbitrary non-commutative ring) Paré and Schelter proved (see [4]) that any matrix $A \in M_n(R)$ satisfies a monic identity in which the leading term is A^k for some integer $k \leq 2^{2^{n-1}}$ and the other summands are of the form $r_0 A r_1 A r_2 \ldots r_{l-1} A r_l$ with $r_0, r_1, \ldots, r_l \in R$ and $0 \leq l \leq k - 1$. An explicit monic identity for 2×2 matrices arising from the argument of [4] and a detailed study of the ideal in $R\langle x \rangle = R * k[x]$ consisting of the polynomials which have as a root the generic $n \times n$ matrix $X = [x_{ij}]$ was given by Robson in [8] ($R = k\langle x_{ij} \rangle$ is the free associative algebra over a field k and $R\langle x \rangle = R * k[x]$ is the free associative k-algebra in one more indeterminate x). Further results in this direction can be found in [5], [6] and [7].

The aim of the present paper is to define the right characteristic polynomial

$$p(x) = \lambda_0 + \lambda_1 x + \cdots + \lambda_{n-1} x^{n-1} + n! x^n$$

in $R[x]$ of an $n \times n$ matrix $A \in M_n(R)$ and to derive a corresponding identity for A (here R is an arbitrary unitary ring). We obtain a Cayley-Hamilton identity with right matrix coefficients of the following form:

$$(\lambda_0 I + C_0) + A(\lambda_1 I + C_1) + \cdots + A^{n-1}(\lambda_{n-1} I + C_{n-1}) + A^n(n!I + C_n) = 0,$$

where $I \in M_n(R)$ is the identity matrix and the entries of the $n \times n$ matrices C_i, $0 \le i \le n$ are in the additive subgroup $[R, R]$ of R generated by the commutators $[x, y] = xy - yx$ with $x, y \in R$ (a more precise description of the entries in the C_i's can be deduced from the proof). Note that a similar identity with left matrix coefficients can be obtained analogously. If R is commutative, then

$$C_0 = C_1 = \cdots = C_{n-1} = C_n = 0$$

and our identity gives the $n!$ times scalar multiple of the classical Cayley-Hamilton identity for A.

We shall make extensive use of the results of [9], in order to provide a self contained treatment, we recall all the necessary prerequisites from [9].

2. The characteristic polynomial. Let R be an arbitrary unitary ring, the *preadjoint* of an $n \times n$ matrix

$$A = [a_{ij}]\ ,\ a_{ij} \in R\ ,\ 1 \le i, j \le n$$

is defined as $A^* = [a^*_{rs}] \in M_n(R)$, where

$$a^*_{rs} = \sum_{\tau,\rho} \operatorname{sgn}(\rho) a_{\tau(1)\rho(\tau(1))} \cdots a_{\tau(s-1)\rho(\tau(s-1))} a_{\tau(s+1)\rho(\tau(s+1))} \cdots a_{\tau(n)\rho(\tau(n))}$$

and the sum is taken over all permutations τ of the set $\{1, \ldots, s-1, s+1, \ldots, n\}$ and ρ of the set $\{1, 2, \ldots, n\}$ with $\rho(s) = r$. The *right determinant* of A is the trace of the product matrix AA^*:

$$\operatorname{r}\det(A) = \operatorname{tr}(AA^*).$$

Our development is based on the following crucial result of [9].

Theorem 2.1. *The product $AA^* \in M_n(R)$ can be written in the following form:*

$$AA^* = b_{11} I + C,$$

where b_{11} is the $(1, 1)$ entry in $AA^* = [b_{ij}]$ and $C = [c_{ij}]$ is an $n \times n$ matrix with $c_{11} = 0$ and each c_{ij}, $1 \le i, j \le n$ is a sum of commutators of the form $[u, v]$ $(u, v \in R)$.

Remark 2.2. The proof of Theorem 2.1 yields that each c_{ij}, $1 \leq i, j \leq n$, $(i, j) \neq (1, 1)$ is a sum of commutators of the form $[\pm a', a'']$, where a' and a'' are products of certain entries of A.

Corollary 2.3. *For the product $AA^* \in M_n(R)$ we have:*

$$nAA^* = \mathrm{tr}(AA^*)I + C',$$

where $C' = [c'_{ij}]$ is an $n \times n$ matrix with $\mathrm{tr}(C') = 0$ and each c'_{ij}, $1 \leq i, j \leq n$ is a sum of commutators of the form $[u, v]$ $(u, v \in R)$.

Proof. The claim easily follows from

$$\begin{aligned} C' &= nAA^* - \mathrm{tr}(AA^*)I \\ &= n(b_{11}I + C) - \mathrm{tr}(AA^*)I \\ &= (nb_{11} - \mathrm{tr}(AA^*))I + nC \\ &= ((b_{11} - b_{11}) + (b_{11} - b_{22}) + \cdots + (b_{11} - b_{nn}))I + nC \\ &= (-c_{22} - \cdots - c_{nn})I + nC. \end{aligned}$$ □

Let $R[x]$ denote the ring of polynomials of the single commuting indeterminate x, with coefficients in R. The *right characteristic* polynomial of A is the right determinant of the $n \times n$ matrix $xI - A$ in $M_n(R[x])$:

$$p(x) = \mathrm{tr}((xI - A)(xI - A)^*) = \lambda_0 + \lambda_1 x + \cdots + \lambda_{d-1}x^{d-1} + \lambda_d x^d \in R[x].$$

Proposition 2.4. *If $p(x) = \lambda_0 + \lambda_1 x + \cdots + \lambda_{d-1}x^{d-1} + \lambda_d x^d$ is the right characteristic polynomial of the $n \times n$ matrix $A \in M_n(R)$ then $d = n$ and $\lambda_d = n!$.*

Proof. Take $(xI - A)^* = [h_{ij}(x)]$ and consider the trace of the product matrix $(xI - A)(xI - A)^*$:

$$\begin{aligned} p(x) &= \sum_{1 \leq i,j \leq n} (xI - A)_{ij} h_{ji}(x) \\ &= (x - a_{11})h_{11}(x) + \cdots + (x - a_{nn})h_{nn}(x) + \sum_{1 \leq i,j \leq n, i \neq j} (-a_{ij})h_{ji}(x). \end{aligned}$$

In view of the definition of the preadjoint, we can see that the degree of $h_{ij}(x)$ is $n - 2$ if $i \neq j$ and the leading monomial of $h_{ii}(x)$ is $(n - 1)!x^{n-1}$. Thus the leading monomial $n!x^n$ of $p(x)$ comes from

$$(x - a_{11})h_{11}(x) + \cdots + (x - a_{nn})h_{nn}(x).$$ □

Proposition 2.5. *If the ring R is commutative, then we have*

$$p(x) = n!\det(xI - A)$$

for the right characteristic polynomial $p(x)$ of the $n \times n$ matrix $A \in M_n(R)$. Thus $p(x)$ is the $n!$ times scalar multiple of the ordinary characteristic polynomial of A.

Proof. Now we have

$$\begin{aligned} p(x) &= \mathrm{tr}((xI - A)(xI - A)^*) = \mathrm{tr}((xI - A)(n-1)!\mathrm{adj}(xI - A)) \\ &= \mathrm{tr}((n-1)!\det(xI - A)I) = n(n-1)!\det(xI - A). \end{aligned}$$

We used the fact that now $(xI - A)^*$ can be expressed as

$$(xI - A)^* = (n-1)!\mathrm{adj}(xI - A)$$

by the ordinary adjoint of $xI - A$ (see also in [9]). □

3. The Cayley-Hamilton identity.

Theorem 3.1. *If $p(x) = \lambda_0 + \lambda_1 x + \cdots + \lambda_{n-1}x^{n-1} + n!x^n$ is the right characteristic polynomial of the $n \times n$ matrix $A \in M_n(R)$, then we can construct $n \times n$ matrices C_i, $0 \le i \le n$ with entries in $[R, R]$ such that*

$$(\lambda_0 I + C_0) + A(\lambda_1 I + C_1) + \cdots + A^{n-1}(\lambda_{n-1} I + C_{n-1}) + A^n(n!I + C_n) = 0.$$

Proof. We use the natural isomorphism between the rings $M_n(R[x])$ and $M_n(R)[x]$. The application of Corollary 2.3 gives that

$$n(xI - A)(B_0 + B_1 x + \cdots + B_{n-1}x^{n-1}) = p(x)I + C'(x),$$

where

$$(xI - A)^* = [h_{ij}(x)] = B_0 + B_1 x + \cdots + B_{n-1}x^{n-1}$$

with $B_0, B_1, \dots, B_{n-1} \in M_n(R)$ (see the proof of Proposition 2.4) and $C'(x)$ is an $n \times n$ matrix with entries in $[R[x], R[x]]$ (and $\mathrm{tr}(C'(x)) = 0$). Since for $f(x) = \sum_{\nu=1}^{s} u_\nu x^\nu$ and $g(x) = \sum_{\mu=1}^{q} v_\mu x^\mu$ in $R[x]$ the commutator

$$[f(x), g(x)] = f(x)g(x) - g(x)f(x) = \sum_{\nu,\mu}[u_\nu x^\nu, v_\mu x^\mu] = \sum_{\nu,\mu}[u_\nu, v_\mu]x^{\nu+\mu}$$

is a polynomial with coefficients in $[R,R]$, we can write that

$$C'(x) = C_0 + C_1 x + \cdots + C_n x^n,$$

where $C_0, C_1, \ldots, C_n$ are $n \times n$ matrices with entries in $[R,R]$. The matching of the coefficients of the powers of x in the above matrix equation gives that

$$\begin{aligned} -nAB_0 &= \lambda_0 I + C_0, \\ nB_0 - nAB_1 &= \lambda_1 I + C_1, \\ &\vdots \\ nB_{n-2} - nAB_{n-1} &= \lambda_{n-1} I + C_{n-1}, \\ nB_{n-1} &= n! I + C_n. \end{aligned}$$

The left multiplication of $nB_{i-1} - nAB_i = \lambda_i I + C_i$ by A^i ($B_{-1} = B_n = 0$) gives the following sequence of matrix equations:

$$\begin{aligned} -nAB_0 &= \lambda_0 I + C_0, \\ nAB_0 - nA^2B_1 &= A\lambda_1 + AC_1, \\ &\vdots \\ nA^{n-1}B_{n-2} - nA^nB_{n-1} &= A^{n-1}\lambda_{n-1} + A^{n-1}C_{n-1}, \\ nA^nB_{n-1} &= A^n n! + A^n C_n. \end{aligned}$$

Thus we obtain that

$$(\lambda_0 I + C_0) + A(\lambda_1 I + C_1) + \cdots + A^{n-1}(\lambda_{n-1} I + C_{n-1}) + A^n(n! I + C_n) =$$

$$= (-nAB_0) + (nAB_0 - nA^2B_1) + \cdots + (nA^{n-1}B_{n-2} - nA^nB_{n-1}) + (nA^nB_{n-1}) = 0. \square$$

In view of the construction of the C_i's in the above proof, it is reasonable to call

$$P(x) = n(xI - A)(xI - A)^* = p(x)I + C_0 + C_1 x + \cdots + C_n x^n$$

the *generalized right characteristic polynomial* of $A \in M_n(R)$. Thus we have

$$np(x) = \mathrm{tr}(P(x)).$$

Proposition 3.2. *If R is an algebra over a field K of characteristic zero and $T \in GL_n(K)$ then we have*

$$p_{TAT^{-1}}(x) = p_A(x) \text{ and } P_{TAT^{-1}}(x) = TP_A(x)T^{-1}$$

for the right characteristic polynomial $p_A(x) \in R[x]$ and the generalized right characteristic polynomial $P_A(x) \in M_n(R)[x]$ of $A \in M_n(R)$.

Proof. In [1] Domokos proved that $(TAT^{-1})^* = TA^*T^{-1}$, whence

$$\begin{aligned} P_{TAT^{-1}}(x) &= n(xI - TAT^{-1})(xI - TAT^{-1})^* \\ &= nT(xI - A)T^{-1}(T(xI - A)T^{-1})^* \\ &= nT(xI - A)T^{-1}T(xI - A)^*T^{-1} \\ &= Tn(xI - A)(xI - A)^*T^{-1} \\ &= TP_A(x)T^{-1} \end{aligned}$$

and

$$np_{TAT^{-1}}(x) = \mathrm{tr}(TP_A(x)T^{-1}) = \mathrm{tr}(P_A(x)) = np_A(x)$$

follows. Since $\dfrac{1}{n} \in K$, we conclude that $p_{TAT^{-1}}(x) = p_A(x)$. □

REFERENCES

[1] M. Domokos. Cayley-Hamilton theorem for 2×2 matrices over the Grassmann algebra. *J. Pure Appl. Algebra* **133** (1998), 69–81.

[2] A. R. Kemer. Varieties of $\mathbb{Z}_2$-graded algebras. *Math. USSR Izv.* **25** (1985), 359–374.

[3] A. R. Kemer. Ideals of Identities of Associative Algebras. Translations of Math. Monographs vol. **87**, 1991, AMS Providence, Rhode Island.

[4] R. Paré, W. Schelter. Finite extensions are integral. *J. Algebra* **53** (1978), 477–479.

[5] K. R. Pearson. A lower bound for the degree of polynomials satisfied by matrices. *J. Aust. Math. Soc., Ser. A* **27** (1979), 430–436.

[6] K. R. Pearson. Degree 7 monic polynomials satisfied by a 3×3 matrix over a noncommutative ring. *Commun. Algebra* **10** (1982), 2043–2073.

[7] K. R. Pearson. The Robson cubics generate all polynomials satisfied by the general 2×2 matrix. *Commun. Algebra* **10** (1982), 2075–2084.

[8] J. C. Robson. Polynomials satisfied by matrices. *J. Algebra* **55** (1978), 509–520.

[9] J. Szigeti. New determinants and the Cayley-Hamilton theorem for matrices over Lie nilpotent rings. *Proc. Amer. Math. Soc.* **125**, *8* (1997), 2245–2254.

[10] J. Szigeti. On the characteristic polynomial of supermatrices. *Israel J. Math.* **107** (1998), 229–235.

Institute of Mathematics
University of Miskolc
Miskolc, Hungary 3515
e-mail: jeno.szigeti@uni-miskolc.hu

Received July 15, 2006