24. On Cayley-Hamilton's Theorem and Amitsur-Levitzki's Identity

By Yoshio AGAOKA
Department of Mathematics, Kyoto University

(Communicated by Shokichi IYANAGA, M. J. A., March 12, 1987)

1. The purpose of this note is to prove a generalization of the classical Cayley-Hamilton's theorem and a tensor version of Amitsur-Levitzki's identity concerning matrices.

Let V be an n-dimensional vector space over the field of complex numbers and $A_1, \cdots, A_p$ be linear endomorphisms of V. We define a linear map $A_1\wedge\cdots\wedge A_p : \wedge^p V \longrightarrow \wedge^p V$ ($\wedge^p V$ is the skew symmetric tensor product of V) by

$$(A_1\wedge\cdots\wedge A_p)(u_1\wedge\cdots\wedge u_p)=(1/p!)\sum_{\sigma\in\mathfrak{S}_p}(-1)^\sigma A_1 u_{\sigma(1)}\wedge\cdots\wedge A_p u_{\sigma(p)},$$

where $(-1)^\sigma$ is the signature of the permutation $\sigma\in\mathfrak{S}_p$ and $u_1, \cdots, u_p \in V$. Note that the equality $A_{\sigma(1)}\wedge\cdots\wedge A_{\sigma(p)}=A_1\wedge\cdots\wedge A_p$ holds for any permutation $\sigma\in\mathfrak{S}_p$. For $X\in\mathrm{End}(V)$, we define invariants $f_i(X)\in \boldsymbol{C}$ by

$$\det(\lambda I-X)=\sum_{i=0}^n f_i(X)\lambda^{n-i},$$

where I is the identity matrix. Then we have

Theorem 1. *Let X be a linear endomorphism of V and p be an integer $(1\leqq p\leqq n)$. Then, by putting $r=n+1-p$, the following identity holds:*

$$(1)\quad \begin{aligned}&\sum_{\substack{a_1+\cdots+a_p=r\\ a_i\geq 0}} X^{a_1}\wedge\cdots\wedge X^{a_p}+f_1(X)\sum_{\substack{a_1+\cdots+a_p=r-1\\ a_i\geq 0}} X^{a_1}\wedge\cdots\wedge X^{a_p}+\cdots\\ &+f_{r-1}(X)\sum_{\substack{a_1+\cdots+a_p=1\\ a_i\geq 0}} X^{a_1}\wedge\cdots\wedge X^{a_p}+f_r(X)\cdot\overbrace{I\wedge\cdots\wedge I}^{p}=0 :\\ &\hspace{20em}\wedge^p V\longrightarrow\wedge^p V,\end{aligned}$$

where the sum is taken over all the combinations of integers $\{a_i\}$ satisfying the conditions under Σ. (We consider $X^0=I$.)

Remark. In the case $p=1$, the above identity is reduced to the form:

$$X^n+f_1(X)X^{n-1}+\cdots+f_n(X)\cdot I=0 : V\longrightarrow V,$$

which is nothing but the classical Cayley-Hamilton's theorem.

Proof. We have only to prove the theorem in case where X is a diagonal matrix because such a matrix constitutes a dense subset of the space of matrices. Let $\{\alpha_1, \cdots, \alpha_n\}$ be the eigenvalues of X and $\{e_1, \cdots, e_n\}$ be a basis of V such that $Xe_i=\alpha_i e_i$. We prove that the element $e_1\wedge\cdots\wedge e_p \in\wedge^p V$ is mapped to 0 by the left hand side of the identity (1). We put $V_1=\{e_1, \cdots, e_p\}$ and $V_2=\{e_{p+1}, \cdots, e_n\}$. First, we have

$$(2)\quad \begin{aligned}(X^{a_1}\wedge\cdots\wedge X^{a_p})(e_1\wedge\cdots\wedge e_p)&=(1/p!)\sum_{\mathfrak{S}_p}(-1)^\sigma X^{a_1}e_{\sigma(1)}\wedge\cdots\wedge X^{a_p}e_{\sigma(p)}\\ &=(1/p!)\sum_{\mathfrak{S}_p}(-1)^\sigma \alpha_{\sigma(1)}^{a_1}\cdots\alpha_{\sigma(p)}^{a_p}e_{\sigma(1)}\wedge\cdots\wedge e_{\sigma(p)}\\ &=(1/p!)\sum_{\mathfrak{S}_p}\alpha_1^{a_{\sigma(1)}}\cdots\alpha_p^{a_{\sigma(p)}}e_1\wedge\cdots\wedge e_p.\end{aligned}$$

We denote by S_λ and T_λ the Schur functions corresponding to the partition $\lambda=(\lambda_1, \cdots, \lambda_s)$ $(\lambda_1\geq\cdots\geq\lambda_s>0)$ with variables $\{\alpha_1, \cdots, \alpha_p\}$ and $\{\alpha_{p+1}, \cdots, \alpha_n\}$, respectively. (See [2], [3], [5]. For example, $S_1=\alpha_1+\cdots+\alpha_p$ and $T_1=\alpha_{p+1}+\cdots+\alpha_n$.) Then for a positive integer k, the Schur function S_k is equal to the trace of the linear map $X^s: S^k(V_1)\to S^k(V_1)$ defined by $X^s(u_1\circ\cdots\circ u_k)=Xu_1\circ\cdots\circ Xu_k$. ($u_1\circ\cdots\circ u_k\in S^k(V_1)$ is the symmetric tensor product of $u_i\in V_1$.) Hence we have

$$S_k=\sum_{i_1\le\cdots\le i_k}\alpha_{i_1}\cdots\alpha_{i_k}=\sum_{a_1+\cdots+a_p=k}\alpha_1^{a_1}\cdots\alpha_p^{a_p}$$
$$=(1/p!)\sum_{\substack{a_1+\cdots+a_p=k\\ \mathfrak{S}_p}}\alpha_1^{a_{\sigma(1)}}\cdots\alpha_p^{a_{\sigma(p)}}.$$

Combining with the equality (2), we have

(3) $$\sum_{a_1+\cdots+a_p=k}(X^{a_1}\wedge\cdots\wedge X^{a_p})(e_1\wedge\cdots\wedge e_p)=S_k\cdot e_1\wedge\cdots\wedge e_p.$$

Next, we calculate the trace of the linear map $X^\wedge:\bigwedge^k V\to\bigwedge^k V$ defined by $X^\wedge(u_1\wedge\cdots\wedge u_k)=Xu_1\wedge\cdots\wedge Xu_k$. Since $\bigwedge^k V$ is a direct sum of $X^\wedge$-invariant subspaces $\bigwedge^l V_1\otimes\bigwedge^{k-l}V_2$ $(l=0,\cdots,k)$, the trace of $X^\wedge$ is

$$\sum_{l=0}^k\Big(\sum_{i_1<\cdots<i_l}\alpha_{i_1}\cdots\alpha_{i_l}\cdot\sum_{j_1<\cdots<j_{k-l}}\alpha_{j_1}\cdots\alpha_{j_{k-l}}\Big)=\sum_{l=0}^k S_{1^l}T_{1^{k-l}},$$

which is, by definition, equal to $\sum_{i_1<\cdots<i_k}\alpha_{i_1}\cdots\alpha_{i_k}=(-1)^k f_k(X)$. Hence, combining with (3), we have

$$f_k(X)\sum_{a_1+\cdots+a_p=r-k}X^{a_1}\wedge\cdots\wedge X^{a_p}(e_1\wedge\cdots\wedge e_p)$$
$$=(-1)^k\sum_{l=0}^k S_{1^l}T_{1^{k-l}}S_{r-k}\cdot e_1\wedge\cdots\wedge e_p.$$

From this equality, it follows that the element $e_1\wedge\cdots\wedge e_p$ is mapped, by the left hand side of (1), to

$$\sum_{q=0}^r(-1)^{r-q}\{S_q-S_1S_{q-1}+S_{11}S_{q-2}-\cdots+(-1)^qS_{1^q}\}T_{1^{r-q}}\cdot e_1\wedge\cdots\wedge e_p.$$

Using Littlewood-Richardson's rule (cf. [3]), we have

$$S_{1^t}S_{q-t}=S_{q-t+1,1^{t-1}}+S_{q-t,1^t},$$

and substituting this equality into the above, we see that it is equal to $(-1)^rT_{1^r}\cdot e_1\wedge\cdots\wedge e_p$. But this is 0 because $r>\dim V_2$. Hence the identity (1) holds. q. e. d.

2. Next, we state and prove a tensor version of Amitsur-Levitzki's identity by using Theorem 1.*) For $A_1,\cdots,A_p\in\mathrm{End}(V)$, we define an endomorphism $A_1\circ\cdots\circ A_p$ of the symmetric tensor space $S^p(V)$ by

$$(A_1\circ\cdots\circ A_p)(u_1\circ\cdots\circ u_p)=(1/p!)\sum_{\sigma\in\mathfrak{S}_p}A_1u_{\sigma(1)}\circ\cdots\circ A_pu_{\sigma(p)}.$$

It is easy to see that the equality $A_{\sigma(1)}\circ\cdots\circ A_{\sigma(p)}=A_1\circ\cdots\circ A_p$ holds for any permutation $\sigma\in\mathfrak{S}_p$.

Theorem 2. *Let $X_1,\cdots,X_{2n}$ be linear endomorphisms of V. Then the following identity holds:*

(4) $$\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^\sigma(X_{\sigma(1)}X_{\sigma(2)})\circ(X_{\sigma(3)}X_{\sigma(4)})\circ\cdots\circ(X_{\sigma(2n-1)}X_{\sigma(2n)})=0$$
$$:S^n(V)\longrightarrow S^n(V).$$

Remark. It is easy to see that the contraction of the linear map $A_1\circ\cdots\circ A_p:S^p(V)\to S^p(V)$ is

$$\sum_{i=1}^p\mathrm{Tr}\,A_i\cdot A_1\circ\cdots\circ\hat{A}_i\circ\cdots\circ A_p$$
$$+\sum_{i\ne j}(A_iA_j)\circ A_1\circ\cdots\circ\hat{A}_i\circ\cdots\circ\hat{A}_j\circ\cdots\circ A_p.$$

Hence, by contracting the above equality (4) $n-1$-times, we obtain a matrix identity

---

*) Dr. K. Kiyohara kindly communicated to the author another proof using the graph theory.

$$\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}X_{\sigma(1)}X_{\sigma(2)}\cdots X_{\sigma(2n)}=0:V\longrightarrow V,$$

which is Amitsur-Levitzki's identity ([1], [6]).

Proof. For $A_1,\cdots,A_p\in\operatorname{End}(V)$, we define linear maps

$$A_1\square\cdots\square A_p:S^p(V)\longrightarrow\wedge^pV \quad\text{and}\quad A_1\triangle\cdots\triangle A_p:\wedge^pV\longrightarrow S^p(V)$$

by

$$(A_1\square\cdots\square A_p)(u_1\circ\cdots\circ u_p)=\frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p}A_1u_{\sigma(1)}\wedge\cdots\wedge A_pu_{\sigma(p)},$$

and

$$(A_1\triangle\cdots\triangle A_p)(u_1\wedge\cdots\wedge u_p)=\frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p}(-1)^{\sigma}A_1u_{\sigma(1)}\circ\cdots\circ A_pu_{\sigma(p)}.$$

(Note that equalities $A_{\sigma(1)}\square\cdots\square A_{\sigma(p)}=(-1)^{\sigma}A_1\square\cdots\square A_p$ and $A_{\sigma(1)}\triangle\cdots\triangle A_{\sigma(p)}=(-1)^{\sigma}A_1\triangle\cdots\triangle A_p$ hold for any $\sigma\in\mathfrak{S}_p$.) Then the following composition formulas hold.

$$(A_1\wedge\cdots\wedge A_p)(B_1\square\cdots\square B_p)=\frac{1}{p!}\sum(-1)^{\sigma}(A_1B_{\sigma(1)})\square\cdots\square(A_pB_{\sigma(p)})$$

$$\begin{aligned}(A_1\triangle\cdots\triangle A_p)(B_1\square\cdots\square B_p)&=\frac{1}{p!}\sum(-1)^{\sigma}(A_1B_{\sigma(1)})\circ\cdots\circ(A_pB_{\sigma(p)})\\&=\frac{1}{p!}\sum(-1)^{\sigma}(A_{\sigma(1)}B_1)\circ\cdots\circ(A_{\sigma(p)}B_p)\end{aligned}$$

$$(A_1\triangle\cdots\triangle A_p)(B_1\wedge\cdots\wedge B_p)=\frac{1}{p!}\sum(A_1B_{\sigma(1)})\triangle\cdots\triangle(A_pB_{\sigma(p)}).$$

Now, we calculate the following sum of linear maps

$$(5)\qquad \sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}(X_{\sigma(n+1)}\triangle\cdots\triangle X_{\sigma(2n)})\cdot(X_{\sigma(1)}\wedge I\wedge\cdots\wedge I)(X_{\sigma(2)}\square\cdots\square X_{\sigma(n)}\square I):S^n(V)\longrightarrow S^n(V)$$

in two ways. First, from the above composition formula, we have

$$\begin{aligned}&(X_{\sigma(1)}\wedge I\wedge\cdots\wedge I)(X_{\sigma(2)}\square\cdots\square X_{\sigma(n)}\square I)\\&=\frac{1}{n}\sum_{i=2}^{n}(-1)^{i}(X_{\sigma(1)}X_{\sigma(i)})\square\cdots\square X_{\sigma(i-1)}\square X_{\sigma(i+1)}\square\cdots\square X_{\sigma(n)}\square I\\&\quad+\frac{1}{n}(-1)^{n-1}X_{\sigma(1)}\square\cdots\square X_{\sigma(n)}.\end{aligned}$$

Hence, by composing with the map $X_{\sigma(n+1)}\triangle\cdots\triangle X_{\sigma(2n)}$, it follows that (5) is equal to

$$\begin{aligned}&\frac{1}{n\cdot n!}\sum_{\sigma\in\mathfrak{S}_{2n}}\sum_{\tau\in\mathfrak{S}_n}\sum_{i=2}^{n}(-1)^{i}(-1)^{\sigma}(-1)^{\tau}(X_{\sigma\tau(n+1)}X_{\sigma(1)}X_{\sigma(i)})\circ\\&\quad(X_{\sigma\tau(n+2)}X_{\sigma(2)})\circ\cdots\circ(X_{\sigma\tau(n+i-1)}X_{\sigma(i-1)})\circ(X_{\sigma\tau(n+i)}X_{\sigma(i+1)})\circ\cdots\circ\\&\quad(X_{\sigma\tau(2n-1)}X_{\sigma(n)})\circ(X_{\sigma\tau(2n)}I)+\frac{1}{n\cdot n!}(-1)^{n-1}\sum_{\sigma\in\mathfrak{S}_{2n}}\sum_{\tau\in\mathfrak{S}_n}(-1)^{\sigma}(-1)^{\tau}\\&\quad\cdot(X_{\sigma\tau(n+1)}X_{\sigma(1)})\circ\cdots\circ(X_{\sigma\tau(2n)}X_{\sigma(n)})\end{aligned}$$

($\tau\in\mathfrak{S}_n$ is considered as a permutation of the letters $\{n+1,\cdots,2n\}$.)

$$\begin{aligned}&=\frac{1}{n}\sum_{\sigma\in\mathfrak{S}_{2n}}\sum_{i=2}^{n}(-1)^{i}(-1)^{\sigma}(X_{\sigma(n+1)}X_{\sigma(1)}X_{\sigma(i)})\circ(X_{\sigma(n+2)}X_{\sigma(2)})\circ\cdots\circ\\&\quad(X_{\sigma(n+i-1)}X_{\sigma(i-1)})(X_{\sigma(n+i)}X_{\sigma(i+1)})\circ\cdots\circ(X_{\sigma(2n-1)}X_{\sigma(n)})\circ(X_{\sigma(2n)}I)\end{aligned}$$

$$+\frac{(-1)^{n-1}}{n}\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}(X_{\sigma(n+1)}X_{\sigma(1)})\circ\cdots\circ(X_{\sigma(2n)}X_{\sigma(n)})$$

$$=\frac{n-1}{n}\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}(X_{\sigma(n+1)}X_{\sigma(1)}X_{\sigma(2)})\circ(X_{\sigma(n+2)}X_{\sigma(3)})\circ\cdots\circ$$

$$(X_{\sigma(2n-1)}X_{\sigma(n)})\circ X_{\sigma(2n)}+\frac{(-1)^{n-1}}{n}\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}(X_{\sigma(n+1)}X_{\sigma(1)})\circ\cdots\circ$$

$$(X_{\sigma(2n)}X_{\sigma(n)}).$$

On the other hand, since $X_{\sigma(1)}\wedge I\wedge\cdots\wedge I=(1/n)\operatorname{Tr}X_{\sigma(1)}\cdot I\wedge\cdots\wedge I$ (the case $p=n$ in Theorem 1), (5) is equal to

$$\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}\frac{1}{n}\operatorname{Tr}X_{\sigma(1)}\cdot(X_{\sigma(n+1)}\triangle\cdots\triangle X_{\sigma(2n)})(X_{\sigma(2)}\square\cdots\square X_{\sigma(n)}\square I)$$

$$=\frac{1}{n\cdot n!}\sum_{\sigma\in\mathfrak{S}_{2n}}\sum_{\tau\in\mathfrak{S}_n}(-1)^{\sigma}(-1)^{\tau}\operatorname{Tr}X_{\sigma(1)}\cdot(X_{\sigma\tau(n+1)}X_{\sigma(2)})\circ\cdots\circ$$

$$(X_{\sigma\tau(2n-1)}X_{\sigma(n)})\circ(X_{\sigma\tau(2n)}I)$$

$$=\frac{1}{n}\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}\operatorname{Tr}X_{\sigma(1)}\cdot(X_{\sigma(n+1)}X_{\sigma(2)})\circ\cdots\circ(X_{\sigma(2n-1)}X_{\sigma(n)})\circ X_{\sigma(2n)}.$$

From these two expressions, we obtain the equality

$$(6)\quad (n-1)\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}(X_{\sigma(n+1)}X_{\sigma(1)}X_{\sigma(2)})\circ(X_{\sigma(n+2)}X_{\sigma(3)})\circ\cdots\circ(X_{\sigma(2n-1)}X_{\sigma(n)})\circ X_{\sigma(2n)}+(-1)^{n-1}\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}(X_{\sigma(n+1)}X_{\sigma(1)})\circ\cdots\circ(X_{\sigma(2n)}X_{\sigma(n)})$$
$$=\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}\operatorname{Tr}X_{\sigma(1)}\cdot(X_{\sigma(n+1)}X_{\sigma(2)})\circ\cdots\circ(X_{\sigma(2n-1)}X_{\sigma(n)})\circ X_{\sigma(2n)}.$$

Next, starting from the composite

$$\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}(X_{\sigma(n+2)}\triangle\cdots\triangle X_{\sigma(2n)}\triangle I)(X_{\sigma(n+1)}\wedge I\wedge\cdots\wedge I)(X_{\sigma(1)}\square\cdots\square X_{\sigma(n)}):S^n(V)\longrightarrow S^n(V),$$

we obtain, in the same way, the equality

$$(7)\quad (n-1)\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}(X_{\sigma(n+2)}X_{\sigma(n+1)}X_{\sigma(1)})\circ(X_{\sigma(n+3)}X_{\sigma(2)})\circ\cdots\circ(X_{\sigma(2n)}X_{\sigma(n-1)})\circ X_{\sigma(n)}+(-1)^{n-1}\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}(X_{\sigma(n+1)}X_{\sigma(1)})\circ\cdots\circ(X_{\sigma(2n)}X_{\sigma(n)})$$
$$=\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^{\sigma}\operatorname{Tr}X_{\sigma(n+1)}\cdot(X_{\sigma(n+2)}X_{\sigma(1)})\circ\cdots\circ(X_{\sigma(2n)}X_{\sigma(n-1)})\circ X_{\sigma(n)}.$$

We rearrange the indices in (6), (7) and add these two equalities. Then, two terms are cancelled and we obtain the desired identity (4), q.e.d.

References

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

[2] N. Iwahori: The Representation Theory of the Symmetric Group and the General Linear Group. Iwanami, Tokyo (1978) (in Japanese).

[3] G. James and A. Kerber: The Representation Theory of the Symmetric Group. Encyclopedia of Math. and its Appl. vol. 16, Addison-Wesley Publ. Company, Reading (1981).

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

[5] I. G. MacDonald: Symmetric Functions and Hall Polynomials. Oxford Univ. Press (1979).

[6] J. P. Razmyslov: Trace identities of full matrix algebras over a field of characteristic zero. Math. USSR Izv., **8**, 727–760 (1974).