LINEAR RELATIONS BETWEEN HIGHER MATRIX COMMUTATORS

NISAR A. KHAN

1. Introduction. Let M_n denote the space of all n-square matrices over an algebraically closed field F. For $A, B \in M_n$, let

$$(1.1)\qquad \begin{aligned} [A, B] &= AB - BA = B_1, \\ [A, B_{i-1}] &= AB_{i-1} - B_{i-1}A = B_i, \qquad i = 2, 3, \ldots, \end{aligned}$$

define the iterated commutators of A and B. Recently several research papers (**1**, **2**, **4**, and **5**) have appeared on these commutators. In (**1**), Kato and Taussky have proved that for $n = 2$ the iterated commutators of A and B satisfy the linear relation

$$(1.2)\qquad B_3 = (\lambda_1 - \lambda_2)^2 B_1,$$

where λ_1 and λ_2 are the eigenvalues of A. This result has been further generalized by Taussky and Wielandt (**5**) in the following form:

$$(1.3)\qquad B_{2N+1} - \delta_1 B_{2N-1} + \delta_2 B_{2N-3} - \ldots + (-1)^N \delta_N B_1 = 0,$$

where $N = n(n-1)/2$ and δ_k is the kth elementary symmetric function of the $(\lambda_i - \lambda_j)^2$, $1 \leqslant i < j \leqslant n$; $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all the distinct eigenvalues of A.

The purpose of this paper is to establish a linear relation between the iterated commutators B_i when A is quite general and is not necessarily similar to a diagonal matrix. The result we establish here contains (1.2) and (1.3) as special cases. Of course, the technique adopted is purely matricial.

For a fixed $A \in M_n$, let T be the linear map of M_n into itself defined by

$$(1.4)\qquad T(X) = AX - XA \qquad \text{for all } X \in M_n.$$

Then

$$T(B) = B_1$$

and

$$T^i(B) = B_i \qquad \text{for } i = 2, 3, \ldots.$$

Let $E_{ij} \in M_n$ be the matrix with 1 in the (i, j) position and 0 elsewhere. With respect to this basis, ordered lexicographically, it may be checked that T has the matrix representation $A \otimes I - I \otimes A'$ where $\otimes$ indicates the Kronecker product and A' is the transpose of A.

Received March 5, 1963.

If J is the Jordan canonical form for A, i.e., for some non-singular matrix P, $P^{-1}AP = J$, then it is easy to check that

$$A \otimes I - I \otimes A' = (P \otimes Q)(J \otimes I - I \otimes J')(P \otimes Q)^{-1},$$

where Q is the inverse of P', so that the minimal polynomial of T coincides with that of $J \otimes I - I \otimes J'$.

2. The main result. Let A have the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ and let $(x - \lambda_i)^{e_{ij}}$, $j = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, r$, be the elementary divisors of A. Let

$$\sum_{j=1}^{n_i} e_{ij} = m_i,$$

the algebraic multiplicity of λ_i, and

$$\sum_{1}^{r} m_i = n,$$

the order of A. We can assume without loss of generality that the eigenvalues are so arranged that $e_{11} \geqslant e_{21} \geqslant \ldots \geqslant e_{r1}$, and the blocks in the Jordan canonical form of A corresponding to each eigenvalue are so arranged that

$$e_{i1} \geqslant e_{i2} \geqslant \ldots \geqslant e_{in_i}, \qquad i = 1, 2, \ldots, r.$$

Assume that the product of iterated commutators is defined as under:

$$B_i B_j = T^i(B) T^j(B) = T^{i+j}(B) = B_{i+j}, \tag{2.1}$$

for all positive integers i and j.

We now state and prove our main result.

THEOREM. *Let the n-square matrix A be as described above and let B be an arbitrary n-square matrix. Define $B_1, B_2, \ldots$ by* (1.1) *and multiplication of B_i, B_j by* (2.1). *Then*

$$B_{2e_{11}-1} \prod_{1 \leqslant i < j \leqslant r} [B_2 - (\lambda_i - \lambda_j)^2 I]^{e_{i1}+e_{j1}-1} = 0, \tag{2.2}$$

provided all the differences $\lambda_i - \lambda_j$, $i \neq j$, are distinct.

3. Two lemmas.

LEMMA 1. *The matrix*

$$B(x) = xI_n \otimes I_m - [(\lambda I_n + U_n) \otimes I_m - I_n \otimes (\mu I_m + U_m')]$$

is equivalent to $I_{m(n-1)} \dotplus [yI_m + U_m']^n$, where U_n is the auxiliary unit matrix with 1 *in the superdiagonal positions and* 0 *elsewhere, and $y = x - \lambda + \mu$.*

Proof. The given matrix is equal to

$$[x - (\lambda - \mu)]I_n \otimes I_m - [U_n \otimes I_m - I_n \otimes U_m'] = B_1(y)$$

say, where $y = x - (\lambda - \mu)$. Then (i) writing $B_1(y)$ in full as an $n \times n$ block matrix with $yI_m + U_m'$ in the main diagonal positions and $-I_m$ in the super-diagonal positions; (ii) multiplying the first row of blocks by $(yI_m + U_m')$ and adding to the second row; (iii) multiplying the second column of blocks by $(yI_m + U_m')$ and adding to the first; (iv) multiplying the second column of blocks by $-I_m$ and then interchanging the second and first columns, we have

$$B_1(y) \sim I_m \dotplus B_2(y),$$

where

$$B_2(y) = \begin{bmatrix} (yI_m + U_m')^2 & -I_m & 0 \ldots \\ 0 & (yI_m + U_m') & -I_m \ldots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}.$$

Suppose after $r - 1$ such steps that we have

$$B_1(y) \sim I_{m(r-1)} \dotplus B_r(y),$$

where

$$B_r(y) = \begin{bmatrix} (yI_m + U_m')^r & -I_m & 0 \ldots \\ 0 & (yI_m + U_m') & -I_m \ldots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}.$$

In $B_r(y)$ we perform the following operations in succession: (i) multiply the second column of blocks by $(yI_m + U_m')^r$ and add to the first; (ii) multiply the first row of blocks by $(yI_m + U_m')$ and add to the second; (iii) multiply the second column by $-I_m$ and then interchange the first and second columns of blocks to have

$$B_r(y) \sim I_m \dotplus B_{r+1}(y).$$

This completes the induction and the lemma is proved.

Lemma 2. *The minimal polynomial of* $(I_n + U_n) \otimes I_m - I_n \otimes (I_m + U_m')$ *is*

$$[x - (\lambda - \mu)]^{m+n-1}.$$

Proof. By Lemma 1,

$$\begin{aligned} xI_n \otimes I_m &- [(\lambda I_n + U_n) \otimes I_m - I_n \otimes (\mu I_m + U_m')] \\ &= yI_n \otimes I_m + I_n \otimes U_m' - U_n \otimes I_m,\ y = x - (\lambda - \mu), \\ &\sim I_{m(n-1)} \dotplus B_n(y), \end{aligned}$$

when $B_n(y) = [yI_m + U_m']^n$. Thus the minimal polynomial of the given matrix is the determinant of $I_{m(n-1)} \dot{+} B_n(y)$ divided by the greatest common divisor of all subdeterminants of $I_{m(n-1)} \dot{+} B_n(y)$ of order $mn - 1$.

$$\det[I_{m(n-1)} \dot{+} B_n(y)] = \det B_n(y) = y^{mn}.$$

It is easy to see that the greatest common divisor of all the subdeterminants of $I_{m(n-1)} \dot{+} B_n(y)$ of order $mn - 1$ is equal to the greatest common divisor of all the subdeterminants of $B_n(y)$ of order $m - 1$. The (i, j)th entry of $B_n(y)$ is either zero or it is of degree $n - (i - j)$ in y. Thus a term in the expansion of

$$\det B_n(y)[r_1, r_2, \ldots, r_p|s_1, s_2, \ldots, s_p],$$
$$1 \leqslant r_1 < r_2 < \ldots < r_p \leqslant m, \quad 1 \leqslant s_1 < s_2 < \ldots < s_p \leqslant m,$$

is either 0 or a non-zero constant times

$$\prod_{\alpha=1}^{p} y^{n-(r_\alpha - s_{\sigma(\alpha)})},$$

where σ is a permutation of $1, 2, \ldots, p$. Hence any non-zero term of

$$\det B_n(y)[r_1, \ldots, r_p|s_1, \ldots, s_p]$$

is of degree γ, where

$$\gamma = \sum_{\alpha=1}^{p} [n - r_\alpha + s_{\sigma(\alpha)}] = np - \sum_{\alpha=1}^{p} r_\alpha + \sum_{\alpha=1}^{p} s_\alpha.$$

This shows that $\det B_n(y)[r_1, \ldots, r_p|s_1, \ldots, s_p]$ is homogeneous of degree γ. This degree is least when $\sum r_\alpha$ is maximal and $\sum s_\alpha$ is minimal, i.e., when $r_1 = m - p + 1$, $r_2 = m - p + 2, \ldots, r_p = m$, and $s_1 = 1$, $s_2 = 2, \ldots, s_p = p$, in which case

$$\gamma = np - \sum_{1}^{p} \alpha + \sum_{1}^{p} (m - \alpha + 1) = (n - m + p)p.$$

Thus, the determinantal divisor of $I_{n(m-1)} \dot{+} B_n(y)$ of order $nm - 1$ is $y^{(m-1)(n-1)}$. Hence the required minimal polynomial is equal to $y^{mn-(n-1)(m-1)} = y^{n+m-1}$, and Lemma 2 is proved.

4. Proof of the theorem. Let J, the Jordan normal form of A, be given by

$$J = \dot{\sum_{s=1}^{r}} \dot{\sum_{j=1}^{ns}} [\lambda_s I_{esj} + U_{esj}],$$

where I_t denotes the identity matrix of order t, and $\dot{\sum}$ indicates direct sum. Then

$$J \otimes I - I \otimes J' = \left[\dot{\sum_{s=1}^{r}} \dot{\sum_{j=1}^{ns}} (\lambda_s I_{esj} + U_{esj})\right] \otimes I - \left[\dot{\sum_{s=1}^{r}} \dot{\sum_{j=1}^{ns}} I_{esj}\right] \otimes J'$$

$$= \text{the direct sum of the following } \sum_{1}^{r} n_i \text{ matrices:}$$

https://doi.org/10.4153/CJM-1964-031-3 Published online by Cambridge University Press

$$(4.1)\qquad (\lambda_s I_{e_{sj}} + U_{e_{sj}}) \otimes I - I_{e_{sj}} \otimes \sum_{t=1}^{r}{}^{\cdot} \sum_{k=1}^{nt}{}^{\cdot} (\lambda_t I_{etk} + U'_{etk}),$$
$$j = 1, 2, \ldots, n_s;\ s = 1, 2, \ldots, r.$$

By a suitable interchange of rows and columns the matrix (4.1), for a fixed s and j, can be transformed to an equivalent matrix

$$\sum_{t=1}^{r}{}^{\cdot} \sum_{k=1}^{nt}{}^{\cdot} [(\lambda_s I_{e_{sj}} + U_{e_{sj}}) \otimes I_{etk} - I_{e_{sj}} \otimes (\lambda_t I_{etk} + U'_{etk})].$$

Thus we have the following:

minimal polynomial of T

= minimal polynomial of its matrix representation $A \otimes I - I \otimes A'$

= minimal polynomial of $J \otimes I - I \otimes J'$

= least common multiple of the minimal polynomials of the

$$\left(\sum_{i=1}^{r} n_i\right)^2$$

matrices

$$(\lambda_s I_{e_{sj}} + U_{e_{sj}}) \otimes I_{etk} - I_{e_{sj}} \otimes (\lambda_t I_{etk} + U_{etk}),$$
$$j = 1, 2, \ldots, n_s,\ k = 1, 2, \ldots, n_t,\ s, t = 1, 2, \ldots, r,$$

by a well-known result on the minimal polynomial of direct sum of matrices (**3**, 151). By Lemma 2, this minimal polynomial is equal to the least common multiple of the following set of minimal polynomials of the above

$$\left(\sum_{1}^{r} n_i\right)^2$$

matrices:

$$(4.2)\quad \{[x - (\lambda_p - \lambda_q)]^{e_{pi}+e_{qj}-1} | 1 \leqslant i \leqslant n_p, 1 \leqslant j \leqslant n_q, 1 \leqslant p, q \leqslant r\}$$
$$= x^{2e_{11}-1} \prod_{1 \leqslant i < j \leqslant r} [x^2 - (\lambda_i - \lambda_j)^2]^{e_{i1}+e_{j1}-1},$$

because of the order of the e_{ij} with respect to magnitude as prescribed earlier, and because of the fact that all $\lambda_i - \lambda_j$, $i \neq j$, are distinct.

Now T satisfies its minimal polynomial and we have

$$T^{2e_{11}-1} \prod_{1 \leqslant i < j \leqslant r} [T^2 - (\lambda_i - \lambda_j)^2 I]^{e_{i1}+e_{j1}-1} = 0.$$

Since $T^i(B) = B_i$ and since we have decided to take B_{i+j} for $B_i B_j$ for all positive integers i and j, the proof of the theorem is complete.

5. Remarks.

(i) Let A be non-derogatory, i.e., corresponding to each eigenvalue there is only one block in the Jordan form. Then, if $(x - \lambda_i)^{m_i}$, $i = 1, 2, \ldots, r$, $m_1 \geqslant m_2 \geqslant \ldots \geqslant m_r$, are the elementary divisors of A, (2.2) reduces to

$$B_{2m_1-1} \prod_{1 \leqslant i < j \leqslant r} [B_2 - (\lambda_i - \lambda_j)^2 I]^{m_i+m_j-1} = 0. \tag{5.1}$$

(ii) Let the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of A be all distinct. Then $m_i = 1$, $i = 1, 2, \ldots, n$, and (5.1) reduces to

$$B \prod_{1 \leqslant i < j \leqslant n} [B_2 - (\lambda_i - \lambda_j)^2 I] = 0, \tag{5.2}$$

which is the same as the result (1.3) of Taussky and Wielandt **(5)**.

(iii) If $n = 2$ in Remark (ii), then we have (1.2), a result due to Kato and Taussky **(1)**.

(iv) Equation (2.2) gives linear relations between the higher matrix commutators B_i only when all the differences $\lambda_i - \lambda_j$, $i \neq j$, are distinct. If some of the $\lambda_i - \lambda_j$, $i \neq j$, are equal, (4.2) is not the minimal polynomial of $A \otimes I - I \otimes A'$, although it certainly is an annihilating polynomial, whence the required linear relation between the B_i will be of lower degree. There does not seem to be a simple method of expressing the minimal polynomial in such a situation. However, given a specific matrix A, we can write down the minimal polynomial of $A \otimes I - I \otimes A'$ by the method indicated in the proof of our theorem. This minimal polynomial will lead to the required relation. For example, if $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are the distinct eigenvalues of a given matrix A, if $\{(x - \lambda_i)^{e_{ij}}, j = 1, 2, \ldots, n_i, i = 1, 2, 3, 4\}$, with the same restrictions on the orders of e_{ij} as indicated in the beginning of Section 2, are the elementary divisors of A, and if $\lambda_1 - \lambda_2 = \lambda_2 - \lambda_3 = \lambda_3 - \lambda_4$, then the linear relation between the B_i is

$$B_{2e_{11}-1}[B_2 - (\lambda_2 - \lambda_4)^2 I]^{e_{21}+e_{41}-1} \prod_{i=2}^{4} [B_2 - (\lambda_1 - \lambda_i)^2 I]^{e_{11}+e_{i1}-1} = 0.$$

References

1. T. Kato and O. Taussky, *Commutators of A and A**, J. Washington Acad. Sci., *46* (1956), 38–40.
2. M. Marcus and N. A. Khan, *On matrix commutators*, Can. J. Math., *12* (1960), 269–277.
3. S. Perlis, *Theory of matrices* (New York, 1952).
4. M. F. Smiley, *Matrix commutators*, Can. J. Math., *13* (1961), 353–355.
5. O. Taussky and H. Wielandt, *Linear relations between higher additive commutators*, Proc. Am. Math. Soc., *13* (1962), 732–735.

Muslim University,
Aligarh, India