MATRIX COMMUTATORS

M. F. SMILEY

Introduction. A classical theorem states that if a square matrix B over an algebraically closed field F commutes with all matrices X over F which commute with a matrix A over F, then B must be a polynomial in A with coefficients in F **(2)**. Recently Marcus and Khan **(1)** generalized this theorem to double commutators. Our purpose is to complete the generalization to commutators of any order.

Let F be an algebraically closed field and let F_n be the ring of all n by n matrices with elements in F. We define $\Delta_Y Z = [Z, Y] = ZY - YZ$ for all Y, Z in F_n.

THEOREM. *Let $A, B \in F_n$ be such that for some positive integer s, $\Delta_A{}^s X = 0$ for X in F_n implies that $\Delta_X{}^s B = 0$. Let the characteristic of F be 0 or at least n. Then B is a polynomial in A with coefficients in F.*

For $s = 1$ we have the classical theorem except for the restriction on the characteristic of F. For $s = 2$ we have the result of Marcus and Khan with a bit more freedom for the characteristic of F. We feel that even for $s = 2$ our proof has interest. We first observe that $s > 1$ is "rather without meaning" for semi-simple matrices and then we use this observation to reduce our theorem to the classical case. Here we call A in F_n semi-simple in case the roots of the minimal polynominal of A are distinct.

1. Some lemmas.

The results of this section will be used in the next section in which we will prove our theorem.

LEMMA 1. *If A is semi-simple in F_n, then $\Delta_A{}^s X = 0$ for some positive integer s only if $\Delta_A X = 0$.*

Proof. We use induction on s. Let $E_k (k = 1, \ldots, q)$ be the principal idempotents of A so that $A = \mu_1 E_1 + \ldots + \mu_q E_q$ with $\mu_k \in F$ $(k = 1, \ldots, q)$. Then each E_k is a polynomial in A with coefficients in F. The Jacobi identity $[Y, [Z, W]] + [Z, [W, Y]] + [W, [Y, Z]] = 0$ for all Y, Z, W in F_n shows that if $E = E_k$ $(k = 1, \ldots, q)$, then $\Delta_A \Delta_E Y - \Delta_E \Delta_A Y = 0$ for all Y in F_n. Now $\Delta_A{}^s X = [\Delta_A{}^{s-1} X, A] = 0$ gives $[\Delta_A{}^{s-1} X, E] = 0$ and hence $\Delta_A{}^{s-1} \Delta_E X = 0$. By our inductive hypothesis, $\Delta_A \Delta_E X = 0$ from which $\Delta_E{}^2 X = 0$ follows at once. But $\Delta_E{}^2 X = 2EXE + XE - EX = 0$ yields $EX = XE$ upon right and left multiplication by E. Thus $\Delta_E X = 0$ for all $E = E_k$ $(k = 1, \ldots, q)$ and consequently $\Delta_A X = 0$, completing our inductive proof of the lemma.

Received May 12th, 1960.

https://doi.org/10.4153/CJM-1961-030-x Published online by Cambridge University Press

An alternative proof of Lemma 1 is suggested by the referee. We may assume that A is a diagonal matrix and use the well-known matrix representation $L = I \otimes A - A \otimes I$ for Δ_A, where $\otimes$ denotes the Kronecker product. But then L is a diagonal matrix so that L and L^s have the same null-space, and this proves Lemma 1.

At this point we introduce the usual matrix units e_{ij} $(i, j = 1, \ldots, k)$ in F_k. The matrix e_{ij} has 1 in the ith row and jth column and zeros elsewhere.

LEMMA 2. *In F_k, let $C = \lambda I_k + e_{21} + e_{32} + \ldots + e_{kk-1}$ with λ in F and $X = e_{11} + 2e_{22} + \ldots + ke_{kk}$. Then $\Delta_C{}^2X = 0$, and for $Y = (C - \lambda)X$, $\Delta_C{}^2Y = 0$.*

Proof. A simple computation shows that $\Delta_C X = XC - CX = C - \lambda I_k$. Since $\Delta_C(C - \lambda)T = (C - \lambda)\Delta_C T$ for all T in F_k, the lemma follows. (The matrices X and Y are special cases of certain matrices used in (1) on pp. 273-274.)

LEMMA 3. *Let C, X, Y be as in Lemma 2 and let $B \in F_k$. Assume that the characteristic of F is 0 or at least k. Then $[B, X] = 0$ implies that B is a diagonal matrix and $[B, X] = [B, Y] = 0$ implies that B is a scalar matrix.*

Proof. With $B = \Sigma b_{ij}e_{ij}$ we find that $BX = \Sigma jb_{ij}e_{ij}$ and $XB = \Sigma ib_{ij}e_{ij}$. Hence $[B, X] = 0$ gives $b_{ij} = 0$ for $i \neq j$ and $i, j = 1, \ldots, k$. With $B = \text{diag}(b_1, \ldots, b_k)$, $YB = b_1e_{21} + 2b_2e_{32} + \ldots + (k-1)b_{k-1}e_{kk-1}$ and $BY = b_2e_{21} + 2b_3e_{32} + \ldots + (k-1)b_ke_{kk-1}$. Hence $[B, Y] = 0$ yields $b_1 = b_2 = \ldots = b_k$ so that B is a scalar matrix.

2. Proof of the theorem. In this section we use the lemmas of § 1 to prove our theorem. Since we shall use the classical result $(s = 1)$ in our proof, we assume that s is at least 2.

We may clearly assume that $A \in F_n$ is in Jordan normal form:

$$A = \text{diag}(C_1, \ldots, C_t) = \text{diag}(J_1, \ldots, J_q)$$

where each C_i $(i = 1, \ldots, t)$ is an n_i by n_i matrix corresponding to an elementary divisor $(x - \lambda_i)^{p_i}$ of A and each J_k is an m_k by m_k matrix with a single characteristic root μ_k and $\mu_k \neq \mu_l$ for $k \neq l$ $(k, l = 1, \ldots, q)$.

Take $X = \text{diag}(1, \ldots, n)$ and use Lemma 2 to obtain $\Delta_A{}^2X = 0$ and hence $\Delta_X{}^sB = 0$. By Lemma 1, since X is semi-simple, $\Delta_XB = 0$ and B must be diagonal by Lemma 3. We write $B = \text{diag}(B_1, \ldots, B_t)$, $X = \text{diag}(X_1, \ldots, X_t)$ conformally with $A = \text{diag}(C_1, \ldots, C_t)$. With $Y = \text{diag}((C_1 - \lambda_1)X_1, \ldots, (C_t - \lambda_t)X_t)$, we have $\Delta_A{}^2Y = 0$ by Lemma 2 and also $\Delta_A{}^2(X + Y) = 0$. Since $X + Y$ is semi-simple, $\Delta_{X+Y}B = \Delta_YB = 0$. By Lemma 3, $B_i = c_iI_{n_i}$ with c_i in F $(i = 1, \ldots, t)$. Now let C_i and C_{i+1} have the same characteristic root λ and let U be an $(n_i + n_{i+1})$-rowed square matrix whose only non-zero element is 1 in the last row and first column. If $Z = \text{diag}(0, U, 0)$ in conformity with $A = \text{diag}(C_1, \ldots, C_t)$, then $ZA = AZ = \lambda Z$ so that

$\Delta_A Z = 0$. Since $X + Z$ is semi-simple, we obtain $\Delta_{X+Z}B = \Delta_Z B = 0$ from which $c_i = c_{i+1}$ follows. Thus if $B = \text{diag}(B_{01}, \ldots, B_{0q})$ in conformity with $A = \text{diag}(J_1, \ldots, J_q)$, then $B_{0k} = d_k I_{m_k}$ with d_k in F $(k = 1, \ldots, q)$. Now if $[W, A] = 0$ it is well known that $W = \text{diag}(W_1, \ldots, W_q)$ in conformity with $A = \text{diag}(J_1, \ldots, J_q)$. A direct proof of this statement goes as follows. Partition W into blocks W_{kl} in conformity with $A = \text{diag}(J_1, \ldots, J_q)$. If $Y = W_{kl}$ with $k \neq l$, then $[W, A] = 0$ gives $(\rho I + C)Y = YD$ with C and D nil-potent and ρ non-zero in F. Thus $Y(R_D - R_C) = \rho Y$ where R_D, L_C denote right and left multiplications by C, D, respectively. Since C and D are nil-potent, so is $R_D - L_C$, and it follows that $\rho^i Y = 0$, $Y = 0$. Now we see that $[W, A] = 0$ for W in F_n implies that $[W, B] = 0$ and we complete the proof of our theorem by an appeal to the classical case.

References

1. M. Marcus and N. A. Khan, *On matrix commutators*, Can. J. Math., *12* (1960), 269–277.
2. J. H. M. Wedderburn, *Lectures on matrices*, Amer. Math. Soc. Colloq. Pub., *17* (New York, 1934).

University of Iowa
and
University of California, Riverside