ON k-COMMUTATIVE MATRICES*

BY

WILLIAM E. ROTH

INTRODUCTION

DEFINITION 1. *If A and B are two $n\times n$ matrices, then the matrix*

$$(1)\qquad B_k = A^kB - \binom{k}{1}A^{k-1}BA + \binom{k}{2}A^{k-2}BA^2 - \cdots + (-1)^kBA^k$$

is the kth commute of A with respect to B.

Evidently if we designate B by B_0, we have in general

$$(2)\qquad B_{i+1} = AB_i - B_iA \qquad (i = 0, 1, 2, \cdots).$$

The matrices B_i, defined by these relations, have significance in the study of the Lie groups of infinitesimal rotations and have been studied by numerous writers. Particular attention is invited to the references I–XVII.† In the present paper we shall study the commutes of a pair of matrices as a part of matric algebra and shall not attempt to interpret the significance the results may have in modern physical theories.

DEFINITION 2. *The matrix A is k-commutative with respect to B, where A and B are $n\times n$ matrices, if the kth commute of A with respect to B is zero, whereas no commute of A with respect to B of index less than k is zero.*

DEFINITION 3. *The matrices A and B of order n are mutually k-commutative, if say A is k-commutative with respect to B and if B is at most k-commutative with respect to A.*

If A and B are commutative in the usual sense, then they are mutually one-commutative. The quasi-commutative matrices defined by McCoy (XV) are mutually two-commutative in the sense defined above.

In §1, we study general properties of the kth commutes of A with respect to B, with and without the restriction that A be k-commutative with respect to B. In §2, we study more particularly the structure of B, where A is assumed to be in the Jordan canonical form and is k-commutative with respect to B. The solution of the equation

$$(3)\qquad AX - XA = \mu X$$

* Presented to the Society, December 28, 1934; received by the editors April 15, 1935.

† Roman numerals will refer to the references listed at the end of this paper.

is taken up in the third section as a special case under the more general equation

$$X_k = \mu^k X, \tag{4}$$

where X_k is the kth commute of A with respect to X, and μ is a scalar constant. The equation (3) was studied by Killing (I) and Weinstein (IX), and equation (4) by Weyl (IV–VIII), and others (X–XV). Finally the results in the preceding sections are applied in the investigation of sets of anticommutating (XVI) and of semi-commutative matrices (XVII).

1. General results on k-commutative matrices

For convenience in deriving results below, we shall employ a procedure given in some detail in an earlier paper by the writer (XXI); a brief résumé thereof will now be given. Let $M=(m_{ij})$ be an $n\times p$ matrix; then M^R is the $1\times np$ matrix obtained from M by placing its second row on the right of the first, its third on the right of the second, and so on. If N is a $q\times r$ matrix, then $M\langle N\rangle=(m_{ij}N)$ is an $nq\times pr$ matrix, namely, the direct product of M and N. The transpose of M will be designated by M^T. Throughout, matrices will be designated by capital letters, and scalar quantities by lower case letters, save that R and T used as exponents indicate the transformations of matrices noted above.

In accordance with these conventions equation (2) is equivalent to the unilateral equation

$$B_{i+1}^R = B_i^R[A] \qquad (i = 0, 1, 2, \cdots), \tag{5}$$

where $[A]$ is the $n^2\times n^2$ matrix $A^T\langle I\rangle - I\langle A\rangle$. The transformation of equation (2) to (5) is reversible. Equation (1) now takes the simple form

$$B_k^R = B^R[A]^k. \tag{6}$$

If A is k-commutative with respect to B, we have, according to Definition 2,

$$B^R[A]^k = 0; \qquad B^R[A]^h \neq 0, \qquad h < k. \tag{7}$$

The two following theorems are obvious results of definitions:

Theorem 1. *If A is k-commutative with respect to B, then all commutes, B_i, of A with respect to B are zero for $i\geqq k$.*

Theorem 2. *If A is k-commutative with respect to B and to C, then A is at most k-commutative with respect to $bB+cC$, where b and c are scalar multipliers.*

We shall now prove

Theorem 3. *If A is k-commutative with respect to B, then every scalar polynomial in A is at most k-commutative with respect to B.*

Let the scalar polynomial in A be

$$f(A) = a_0 I + a_1 A + a_2 A^2 + \cdots + a_t A^t.$$

The proof of this theorem consists in showing that

$$B^R[f(A)]^k = 0, \tag{8}$$

if (7) is satisfied. Obviously the transpose of $f(A)$ is $f(A^T)$ and

$$[f(A)] = f(A^T)\langle I\rangle - I\langle f(A)\rangle = \sum_{i=1}^{t} a_i[A^i].$$

However

$$\begin{aligned}[A^i] &= [A]\{(A^T)^{i-1}\langle I\rangle + (A^T)^{i-2}\langle A\rangle + \cdots + I\langle A^{i-1}\rangle\},\\ &= \{(A^T)^{i-1}\langle I\rangle + (A^T)^{i-2}\langle A\rangle + \cdots + I\langle A^{i-1}\rangle\}[A],\end{aligned}$$

and we may therefore write

$$[f(A)] = [A]Q,$$

where Q is an $n^2 \times n^2$ matrix and is commutative with $[A]$. Hence

$$[f(A)]^k = [A]^k Q^k.$$

Multiply this equation on the left by B^R and (8) follows because of (7).

THEOREM 4. *If A is k-commutative with respect to B_0 and if B_i is the ith commute of A with respect to B_0, then the commutes B_i $(i=0, 1, 2, \cdots, k-1)$ are linearly independent matrices.*

Suppose that scalar constants α_i $(i=0, 1, \cdots, k-1)$, not all zero, exist such that

$$\alpha_0 B_0 + \alpha_1 B_1 + \cdots + \alpha_{k-1} B_{k-1} = 0;$$

then, according to (6), we have

$$B_0^R\{\alpha_0 I\langle I\rangle + \alpha_1[A] + \cdots + \alpha_{k-1}[A]^{k-1}\} = 0. \tag{9}$$

Multiply the latter on the right by $[A]^{k-1}$ and, according to (7),

$$\alpha_0 B_0^R[A]^{k-1} = \alpha_0 B_{k-1}^R = 0.$$

However, by definition of k-commutative matrices, $B_{k-1} \neq 0$, hence $\alpha_0 = 0$. Similarly, if (9) be multiplied on the right by $[A]^{k-2}$, we find that α_1 must also be zero, and so on. This leads to a contradiction of the assumption that not all α_i are zeros; the theorem is therefore proved.

THEOREM 5. *If A is k-commutative with respect to B and if the degree of no elementary divisor of $A - \lambda I$ exceeds α, then $k \leqq 2\alpha - 1$.*

The matrix $[A]-\lambda I\langle I\rangle$ has at least one elementary divisor $\lambda^{2\alpha-1}$ and none of higher degree (XXI, Theorem 2), if that of highest degree of $A-\lambda I$ is $(a-\lambda)^\alpha$. Therefore $[A]$ satisfies the minimal equation

$$a_0[A]^{2\alpha-1} + a_1[A]^{2\alpha} + \cdots + a_{g-2\alpha+1}[A]^g = 0, \qquad g \geqq 2\alpha - 1,$$

where a_0 is not zero. Multiply this equation on the left by B^R, and by (6) we conclude that

$$a_0 B_{2\alpha-1} + a_1 B_{2\alpha} + \cdots + a_{h-2\alpha+1} B_h = 0,$$

where h is the lesser of the two integers g and $k-1$. If k exceeds $2\alpha-1$, this linear dependence between the commutes B_i $(i=2\alpha-1, 2\alpha, \cdots, h)$ of A with respect to B cannot hold because of Theorem 4. Hence $k\leqq 2\alpha-1$.

Corollary 1. *If $A-\lambda I$ has no elementary divisor whose degree exceeds α and if B_h, the hth commute of A with respect to B, is not zero for $h>2\alpha-1$, then A is k-commutative with respect to B for no finite value of k.*

This corollary follows at once from the theorem above. We may remark that A is k-commutative with respect to no non-zero X satisfying equation (3), but on the other hand every such solution is two-commutative with respect to A and non-zero solutions of this equation may exist; we therefore can conclude that matrices B, such that A is k-commutative with respect to B for no finite k, do exist.

Corollary 2. *There exist no matrices A and B of order less than $(k+1)/2$ such that A is k-commutative with respect to B.*

The degree of the elementary divisor of highest degree of $A-\lambda I$ cannot exceed n. Hence by the theorem above, $k\leqq 2n-1$ in order that A be k-commutative with respect to B. The corollary is proved. McCoy (XV, p. 335) gave a more restrictive result than that of the present corollary in case A and B are mutually two-commutative; namely, that none of second order exist. However, second-order matrices exist such that A is two-commutative with respect to B, and B is not two-commutative with respect to A. Example:

$$A = \begin{pmatrix} 0, & 1 \\ 0, & 0 \end{pmatrix} \qquad B = \begin{pmatrix} a, & b \\ 0, & c \end{pmatrix}, \qquad \text{where } a \neq c.$$

Corollary 3. *If $|A-\lambda I| = (a-\lambda)^n$ and if the degree of no elementary divisor of $A-\lambda I$ exceeds α, then A is k-commutative with respect to every matrix, B, of order n, and for any given B, $k\leqq 2\alpha-1$.*

Weyl (VI, p. 100) originally gave this result. Under the hypotheses of the present corollary $g=2\alpha-1$ and $[A]^{2\alpha-1}=0$ because $[A]-\lambda I\langle I\rangle$ has elementary divisors only of the form λ^h where that of highest degree is $\lambda^{2\alpha-1}$ (XX or XXI Theorem 2). Hence the $(2\alpha-1)$st commute of A with respect to B is zero.

An alternative statement of Theorem 5 is given by

COROLLARY 4. *If A is k-commutative with respect to B, and if the minimal polynomial satisfied by $[A]$ is $\lambda^\beta\phi(\lambda)$, where $\phi(0)\neq 0$, then $k\leqq\beta$.*

Heretofore we have considered the kth commute of A with respect to B only for positive values of k; however, in certain cases Definition 1 may have sense for negative indices as well. Thus the general solution, if it exists, of the equation

$$X_1 = AX - XA = B$$

may be regarded as the (-1)st commute of A with respect to B; and the general solution, X, of the equation

$$X_i = B, \qquad i \geqq 1,$$

where X_i is the ith commute of A with respect to X, is the $(-i)$th commute of A with respect to B. The latter equation is equivalent to

$$X^R[A]^i = B^R;$$

if X, satisfying this equation, exists, it is not unique in that the number of linearly independent solutions is n^2-r_i, where r_i is the rank of $[A]^i$. Hence according to the well known theory of linear non-homogeneous equations the following theorem holds:

THEOREM 6. *If A and B are given matrices of order n, then the $(-i)$th commute of A with respect to B exists, $i>0$, if and only if the matrices*

$$[A]^i \quad \text{and} \quad \begin{pmatrix} B^R \\ [A]^i \end{pmatrix}$$

have the same rank, r_i, and the number of linearly independent $(-i)$th commutes of A with respect to B, $i>0$, is n^2-r_i.

THEOREM 7. *If A is k-commutative with respect to B and if B_i is the ith commute of A with respect to B, then*

$$\begin{aligned} f(A)B &= Bf(A) + B_1f'(A) + \frac{1}{2!}B_2f''(A) + \cdots + \frac{1}{(k-1)!}B_{k-1}f^{(k-1)}(A), \\ Bf(A) &= f(A)B - f'(A)B_1 + \frac{1}{2!}f''(A)B_2 - \cdots + \frac{(-1)^{k-1}}{(k-1)!}f^{(k-1)}(A)B_{k-1}, \end{aligned} \tag{10}$$

where $f(\lambda)$ is a scalar polynomial in λ and $f^{(i)}(\lambda)$ its ith derivative with respect to λ.

Obviously

$$A^T\langle I\rangle = [A] + I\langle A\rangle,$$

and

$$(A^T\langle I\rangle)^r = (A^r)^T\langle I\rangle = ([A] + I\langle A\rangle)^r;$$

but $I\langle A\rangle$ and $[A]$ are commutative matrices and the right member above may therefore be expanded according to the binomial theorem. Multiply the result on the left by B^R; then

$$B^R\{(A^r)^T\langle I\rangle\} = B^R I\langle A^r\rangle + \binom{r}{1} B_1^R I\langle A^{r-1}\rangle + \cdots + B_r^R$$

or

$$A^rB = BA^r + \binom{r}{1} B_1A^{r-1} + \binom{r}{2} B_2A^{r-2} + \cdots + B_r.$$

This relation is equivalent to that derived by Campbell (III, §2), and from it the first identity of the theorem above follows at once. Similarly on the basis of $I\langle A\rangle = A^T\langle I\rangle - [A]$, we can readily prove the second also. The theorem can be generalized to apply for more general functions $f(\lambda)$, and if A is not assumed to be k-commutative with respect to B the formulas still hold save that the right members will not stop with the kth term.

If $A = (a_{ij})$ is an $n\times n$ matrix whose elements a_{ij} $(i, j = 1, 2, \cdots, n)$ are differentiable functions of t, we have

$$\begin{aligned}\frac{dA^r}{dt} &= \binom{r}{1} A_1A^{r-1} + \binom{r}{2} A_2A^{r-2} + \cdots + \binom{r}{k} A_kA^{r-k},\\ &= \binom{r}{1} A^{r-1}A_1 - \binom{r}{2} A^{r-2}A_2 + \cdots + (-1)^{k-1}\binom{r}{k} A^{r-k}A_k,\end{aligned}$$

where $A_1 = (da_{ij}/dt)$, where A_i $(i = 2, 3, \cdots, k)$ is the $(i-1)$st commute of A with respect to its derivative, A_1, and where A is k-commutative with respect to A_1. These formulas may readily be established by mathematical induction. In case A is commutative with its derivative, the right members reduce to the usual result for scalar quantities.

If $f(A)$ is a scalar polynomial (or convergent power series) in A, we readily obtain the following identities:

$$\text{(11)} \qquad \frac{df(A)}{dt} = A_1f'(A) + \frac{1}{2!}A_2f''(A) + \cdots + \frac{1}{k!}A_kf^{(k)}(A),$$

$$= f'(A)A_1 - \frac{1}{2!}f''(A)A_2 + \cdots + \frac{(-1)^{k-1}}{k!}f^{(k)}(A)A_k,$$

where A is k-commutative with respect to its derivative.

THEOREM 8. *If A is $(k+1)$-commutative with respect to X and if the first commute of A with respect to X is equal to the derivative, $A_1=(da_{ij}/dt)$, of A, then*

$$\frac{d}{dt}f(A) = f(A)X - Xf(A),$$

where $f(\lambda)$ is a function of λ such that $f(A)$ converges for all values of t in the interval under consideration.

By hypothesis,

$$A_1 = AX - XA.$$

If in the first formula (11) we add $Xf(A)-Xf(A)$ to the right member and compare the result with (10) we have the result of the theorem above. The restrictions that $f(A)$ be a polynomial in A and that A be k-commutative with respect to its derivative, A_1, may be removed provided proper bounds may be placed upon the elements of A to insure the convergence of $f(A)$.

2. MORE EXPLICIT FORM OF B

We shall now derive restrictions upon the form of B, where that of A is known and where A is k-commutative with respect to B. In the present section and hereafter we shall discontinue the use of subscripts to indicate the commutes of a matrix pair unless the contrary is specifically stated.

THEOREM 9. *If*

$$A = A_1 \dotplus A_2 \dotplus \cdots \dotplus A_r,$$

where the $m_i \times m_i$ matrix A_i $(i=1, 2, \cdots, r)$ has a unique characteristic value a_i and $a_i \neq a_j$, if $i \neq j$, and if A is k-commutative with respect to $B=(B_{ij})$, where B_{ij} $(i, j=1, 2, \cdots, r)$ are $m_i \times m_j$ matrices, then

$$B = B_{11} \dotplus B_{22} \dotplus \cdots \dotplus B_{rr},$$

*and A_i is at most k-commutative with respect to B_{ii} $(i=1, 2, \cdots, r)$.**

It is no restriction to assume that A has the form given above, for by a suitable non-singular transformation it can be brought into this form. Since

* A matrix $M=(M_{ij})$, where M_{ij} are $m_i \times m_j$ matrices and where all $M_{ij}=0$, if $i \neq j$, is here and in the following pages denoted by the notation $M=M_{11} \dotplus M_{22} \dotplus \cdots \dotplus M_{tt}$. A single subscript on the matrices M_{ii} is sufficient in many cases.

A is k-commutative with respect to B, the matrix (1) must be zero and we consequently have the r^2 equations

$$A_i^k B_{ij} - \binom{k}{1} A_i^{k-1} B_{ij} A_j + \binom{k}{2} A_i^{k-2} B_{ij} A_j^2 - \cdots + (-1)^k B_{ij} A_j^k = 0$$

$(i, j=1, 2, \cdots, r)$. These equations must be satisfied by the matrices B_{ij} independently. In the unilateral form they become

$$B_{ij}^{R}[A_i, A_j]^k = 0 \qquad (i, j = 1, 2, \cdots, r), \tag{12}$$

where

$$[A_i, A_j] = A_i^T \langle I_j \rangle - I_i \langle A_j \rangle,$$

and I_α are $m_\alpha \times m_\alpha$ unit matrices. Each of the r^2 equations (12) is equivalent to a system of $m_i m_j$ linear homogeneous equations in the $m_i m_j$ elements of B_{ij}, the matrix of whose coefficients is $[A_i, A_j]^k$. This matrix is singular if and only if $a_i = a_j$ (XVIII–XXI). Therefore $B_{ij}=0$, if $i \neq j$, and in case $i=j$ we see by (12) that A_i is at most k-commutative with respect to B_{ii} $(i=1, 2, \cdots, r)$. This well known result concerning matrices which are commutative in the ordinary sense holds as well for k-commutative matrices. The following theorem is still more precise in defining the structure of B.

THEOREM 10. *If*

$$A = A_1 \dotplus A_2 \dotplus \cdots \dotplus A_s,$$

where $A_i = a_i I_i + D_i$ and I_i and D_i are respectively the unit and the auxiliary unit matrices of order n_i, and if A is k-commutative with respect to $B=(B_{ij})$, where B_{ij} are $n_i \times n_j$ $(i, j, =1, 2, \cdots, s)$ matrices, then $B_{ij}=0$, if $a_i \neq a_j$, and if $a_i = a_j$, B_{ij} has zero elements in at least the first $\{n_i, n_j\} - k$ diagonals, where $\{n_i, n_j\}$ is the greater of the integers n_i and n_j.*†

As in the proof of Theorem 9, we have

$$B_{ij}^{R}[A_i, A_j]^k = 0 \qquad (i, j = 1, 2, \cdots, s),$$

and $B_{ij}=0$, if $a_i \neq a_j$. However, in case $a_i = a_j$,

$$[A_i, A_j] = [D_i, D_j].$$

Hence

* The auxiliary unit matrices D_i of order n_i are here understood to have $n_i - 1$ unit elements in the first diagonal above the principal diagonal and to have zero elements elsewhere.

† Diagonals are here numbered consecutively beginning with that containing the lower left element of the blocks B_{ij}.

$$B^{R}_{ij}[D_i, D_j]^k = 0 \qquad (i, j = 1, 2, \cdots, s), \tag{13}$$

or

$$D_i^k B_{ij} - \binom{k}{1} D_i^{k-1} B_{ij} D_j + \binom{k}{2} D_i^{k-2} B_{ij} D_j^2 - \cdots + (-1)^k B_{ij} D_j^k = 0.$$

Let

$$B_{ij} = \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix},$$

where C_1, C_2, C_3, C_4 are respectively $\alpha \times (n_j - \beta)$, $\alpha \times \beta$, $(n_i - \alpha) \times (n_j - \beta)$, $(n_i - \alpha) \times \beta$ matrices; then

$$D_i^\alpha B_{ij} D_j^\beta = \begin{pmatrix} 0 & C_3 \\ 0 & 0 \end{pmatrix}.$$

Because of this fact we can conclude that B_{ij}, which satisfies (13), must have only zero elements in at least the first $\{n_i, n_j\} - k$ diagonals where $\{n_i, n_j\}$ is the greater of the integers n_i and n_j. This proves the theorem.

However, in case $n_i = n_j$ and $a_i = a_j$, we can show that the elements in the $(n_i - k + 1)$st diagonal of B_{ij} are likewise zeros provided $k > 1$, since in this case these elements must satisfy linear homogeneous equations with non-zero determinants. This fact, together with the form of B as demonstrated above, leads us to the following theorem:

THEOREM 11. *If $A - \lambda I$ has the elementary divisors $(a_i - \lambda)^{n_i}$ $(i = 1, 2, \cdots, s)$, if A is two-commutative with respect to B, and if $n_i \neq n_j \pm 1$ in case $a_i = a_j$, then the characteristic values of $f(A, B)$, where $f(\lambda, \mu)$ is a scalar polynomial in λ and μ, are in the set $f(a_i, b_h)$ where b_h $(h = 1, 2, \cdots, t)$ are the distinct characteristic values of B.*

Under the hypotheses of this theorem we add no restrictions upon A and B if we assume that A is in the Jordan canonical form given in Theorem 10. The matrix B will be an umbral matrix (XXII), whose blocks B_{ij} are zero in case $a_i \neq a_j$, and therefore with A has the property stated in the theorem above, which we shall designate as the property P.

In case A and B are mutually two-commutative, McCoy (XV, Theorem 5) shows that the third hypothesis of the theorem above may be omitted. The property P does not carry over to mutually k-commutative matrices, where k exceeds 2. For example, the matrices

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} a & 0 & b & 0 \\ 3e & a & 0 & -b \\ 0 & 4e & a & 0 \\ 0 & 0 & 3e & a \end{pmatrix}, \qquad e \neq 0,$$

are mutually three-commutative and the characteristic values of $A+B$ are not those of B. Therefore ordinary commutative matrices and the quasi-commutative matrices of McCoy are the only types of mutually k-commutative matrices which necessarily have the property P.

3. The equation $X_k = \mu^k X$

Evidently every matrix X which satisfies the equation

$$(3) \qquad AX - XA = \mu X$$

will likewise satisfy the equation

$$(4) \qquad X_k = \mu^k X,$$

where X_k is the kth commute of A with respect to X and μ is a non-zero scalar constant. On the other hand not all solutions of (4) satisfy (3). We shall confine our attention to (4).

We may without restrictions upon the problem assume that A is in the Jordan canonical form

$$A = A_1 \dotplus A_2 \dotplus \cdots \dotplus A_s,$$

where $A_i = a_i I_i + D_i$ and I_i and D_i are respectively the unit and the auxiliary unit matrices of order n_i. Under these assumptions the elementary divisors of $A - \lambda I$ are $(a_i - \lambda)^{n_i}$ $(i = 1, 2, \cdots, s)$. Let $X = (X_{ij})$, where X_{ij} $(i, j = 1, 2, \cdots, s)$ are $n_i \times n_j$ matrices; then (4) is equivalent to the s^2 equations

$$(14) \qquad X_{ij}^R \{ [A_i, A_j]^k - \mu^k I_i \langle I_j \rangle \} = 0 \quad (i, j = 1, 2, \cdots, s).$$

The necessary and sufficient condition that X_{ij} be a non-zero matrix is that the $n_i n_j \times n_i n_j$ matrix

$$(15) \qquad [A_i, A_j]^k - \mu^k I_i \langle I_j \rangle \qquad (i, j = 1, 2, \cdots, s)$$

be singular. It has the characteristic value $(a_i - a_j)^k - \mu^k$ repeated $n_i n_j$ times (XVIII–XXI). Hence the necessary and sufficient condition that X_{ij} be a non-zero matrix is that $(a_i - a_j)^k - \mu^k = 0$. Moreover, since $\mu \neq 0$, we have $X_{ii} = 0$ $(i = 1, 2, \cdots, s)$; that is, the trace of X, any solution of (3), is zero (compare IV). These properties of X are invariants under the usual transformations of matrices to normal form. Hence we have the theorems below.

THEOREM 12. *The necessary and sufficient condition that equation* (4) *have a solution other than the trivial solution* $X=0$ *is that* A *have at least two characteristic values* a *and* b *such that* $(a-b)^k=\mu^k$.

This result was obtained by Weinstein (IX) for the case $k=1$.

THEOREM 13. *The trace of every solution*, X, *of* (4) *is zero.*

We shall now prove

THEOREM 14. *If* $A-\lambda I$ *has the elementary divisors* $(a_i-\lambda)^{n_i}$ $(i=1,2,\cdots,s)$, *and if* X *is a solution of* (4), *where* k *is an odd integer, then* X *is a nil-potent matrix if it is possible so to arrange the characteristic values* a_i *of* A *that*

$$\left(\frac{a_i-a_j}{\mu}\right)^k \neq 1 \quad \textit{for} \quad i>j.$$

In this case all X_{ij}, $i\geqq j$, are zero, and all non-zero X_{ij}, if such exist, lie above the principal diagonal of X. That is, X is a nil-potent matrix.

We shall now expose the exact form of X_{ij} in case $(a_i-a_j)^k/\mu^k=1$. The matrix (15) in this case becomes

$$\begin{aligned}&\left\{(a_i-a_j)^k I_i\langle I_j\rangle - N\right\}^k - \mu^k I_i\langle I_j\rangle\\ &\quad = \binom{k}{1}(a_i-a_j)^{k-1}N + \binom{k}{2}(a_i-a_j)^{k-2}N^2+\cdots+N^k,\end{aligned}$$

where $N=[D_i, D_j]$. Let the right member be given by NQ; then Q is a non-singular matrix since N is nil-potent. The equation (14) consequently becomes

$$\overset{R}{X_{ij}}N = 0,$$

or

$$D_iX_{ij} - X_{ij}D_j = 0. \tag{16}$$

This is the well known relation which arises in the study of matrices X commutative with the Jordan canonical matrix A, save that in the present case (16) holds if $(a_i-a_j)^k=\mu^k$, and not if $a_i=a_j$ as in case A and X are commutative. Therefore $X=(X_{ij})$ $(i, j=1, 2, \cdots, s)$, a solution of (4), is such that in case $(a_i-a_j)^k=\mu^k$, X_{ij} has zero elements in the first $\{n_i, n_j\}-1$ diagonals, and the elements in each of the remaining diagonals of X_{ij} are all equal but arbitrary and independent of those of another diagonal. If $(a_i-a_j)^k\neq\mu^k$, then $X_{ij}=0$. From the structure of X here discussed, the following theorem is at once evident, since if it is not satisfied then X will have at least one row or column of zero elements and will be singular.

THEOREM 15. *If $A-\lambda I$ has the elementary divisors $(a_i-\lambda)^{n_i}$ $(i=1,2,\cdots,s)$, then the necessary and sufficient condition that the equation* (4) *have a non-singular solution X is that for every i $(i=1,2,\cdots,s)$ there exist at least one j $(j=1,2,\cdots,s)$, and for every j there exist at least one i, such that*

$$n_i = n_j \quad \text{and} \quad (a_i - a_j)^k = \mu^k \quad (i, j = 1, 2, \cdots, s).$$

If the matrix A is in the Jordan canonical form, then X, a solution of (4), is an umbral matrix (XXII, Definition 3) and consequently

$$|X| = |X_\alpha| \cdot |X_\beta| \cdot \cdots \cdot |X_\rho|,$$

where $X_h=(X_{ij})$ $(h=\alpha,\beta,\cdots,\rho)$ and i,j run over only those values for which $n_i=n_j=n_h$, and $n_\alpha, n_\beta, \cdots, n_\rho$ are the distinct values of n_i $(i=1,2,\cdots,s)$. (See XXII, Theorem III.) This fact makes the restriction $n_i=n_j$ and $(a_i-a_j)^k=\mu^k$ a necessary one, else $|X|=0$.

4. SETS OF SEMI-COMMUTATIVE MATRICES

If A and B satisfy the relations

$$AB = \omega BA \quad \text{and} \quad A^k = B^k = I, \tag{17}$$

where ω is a primitive kth root of unity, they have been called semi-commutative by Williamson (XVII). On the basis of the first equation (17) we can readily show that

$$B_i = (\omega - 1)^i BA^i \qquad (i = 1, 2, \cdots),$$

where B_i is the ith commute of A with respect to B. Consequently, because of the second restriction upon A in (17), we have

$$B_k = (\omega - 1)^k B. \tag{18}$$

This proves

THEOREM 16. *If A is a member of the set of semi-commutative matrices, then a second member of that set is a solution of the equation*

$$X_k = (\omega - 1)^k X,$$

where X_k is the kth commute of A with respect to X.

The theory developed in §3 is applicable in this section; however, the results there obtained are more general than necessary in the present case. A special study of (17) is superfluous in view of Williamson's results (XVII).

Bibliography

I. Killing, Mathematische Annalen, vol. 31 (1888), pp. 252–290.
II. Campbell, Proceedings of the London Mathematical Society, vol. 28 (1897), pp. 381–390.
III. ——— Ibid., vol. 29 (1898), pp. 14–32.
IV. Weyl, Mathematische Zeitschrift, vol. 12 (1922), pp. 114–146.
V. ——— Ibid., vol. 17 (1923), pp. 293–320.
VI. ——— *Mathematische Analyse des Raumproblems*, Springer, 1923.
VII. ——— *Raum, Zeit, Materie*, 4te Auflage, Springer, 1921.
VIII. ——— *The Theory of Groups and Quantum Mechanics*, Dutton, 1931, p. 272.
IX. Weinstein, Mathematische Zeitschrift, vol. 16 (1923), pp. 78–91.
X. Buhl, *Aperçus Modernes sur la Théorie des Groupes Continus et Finis*, Mémorial des Sciences Mathématiques, fascicule 33, 1928. (Includes bibliography.)
XI. ——— *Gravifiques, Groupes, Mécaniques*, ibid., fascicule 62, 1934. (Includes bibliography.)
XII. van der Waerden, Mathematische Zeitschrift, vol. 36 (1933), pp. 780–786.
XIII. ——— Ibid., vol. 37 (1933), pp. 446–462.
XIV. Casimir und van der Waerden, Mathematische Annalen, vol. 111 (1935), pp. 1–12.
XV. McCoy, Transactions of the American Mathematical Society, vol. 36 (1934), pp. 327–340.
XVI. Eddington, Journal of the London Mathematical Society, vol. 7 (1933), pp. 58–69.
XVII. Williamson, Proceedings of the Edinburgh Mathematical Society, (2), vol. 3 (1933), pp. 179–188, 231–240.

References to matrix theory

XVIII. Stéphanos, Journal de Mathématiques, (5), vol. 6 (1900), pp. 73–128.
XIX. MacDuffee, *The Theory of Matrices*, Springer, 1933, Theorem 43.8.
XX. Aitken, Proceedings of the Edinburgh Mathematical Society, (2), vol. 1 (1928), pp. 135–138.
XXI. Roth, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 461–468.
XXII. ——— These Transactions, vol. 39 (1936), pp. 234–243.

University of Wisconsin, Extension Division,
Milwaukee, Wis.