ON THE CHARACTERISTIC VALUES OF THE MATRIX $f(A, B)$*

BY

WILLIAM E. ROTH

DEFINITION 1. *If A and B, two matrices of order n, have the characteristic values a_i $(i=1, 2, \cdots, n)$ and b_i $(i=1, 2, \cdots, n)$, respectively, then A and B have the property P if and only if every scalar polynomial $f(A, B)$ in A and B has the characteristic values $f(a_i, b_i)$ $(i=1, 2, \cdots, n)$; and the characteristic value a_i of A is said to be associated with b_i of B if $f(a_i, b_i)$ is a characteristic value of every $f(A, B)$.*

It is well known that pairs of commutative matrices have the property P† and recently McCoy‡ showed that quasi-commutative matrices likewise have this property. In §1 we prove a general condition which is necessarily satisfied by a matrix pair having the property P; moreover we show how the characteristic values of two such matrices must be paired or associated. When A is assumed to be a Jordan canonical matrix, we display in §2 the form of B such that A and B have the property P. The results obtained include, as far as is known to the author, all known matrix pairs having the property P,§ and include as well those obtained by Bruton.‖

1. NECESSARY CONDITIONS

THEOREM I. *If*

$$A = A_1 \dot{+} A_2 \dot{+} \cdots \dot{+} A_r,¶$$

* Presented to the Society, April 19, 1935; received by the editors February 25, 1935.

After the present paper had been sent to the editors, I learned of Professor Williamson's article, *The simultaneous reduction of two matrices to triangle form*, American Journal of Mathematics, vol. 57 (April, 1935), pp. 281–293. Professor Williamson's very interesting paper gives necessary and sufficient conditions for the existence of the property P in case one of the matrices, say A, is such that its characteristic equation is also its minimal equation, whereas the present article does not make this restriction on A and does not establish conditions which are both necessary and sufficient.

† Frobenius, *Über vertauschbare Matrizen*, Berliner Sitzungsberichte, 1896, pp. 601–614.

‡ McCoy, *On quasi-commutative matrices*, these Transactions, vol. 36 (1934), pp. 327–340.

§ MacDuffee, *The Theory of Matrices*, Berlin, 1933, in §16 gives a concise résumé of known results.

‖ Bruton, *Certain aspects of the theory of equations of a pair of matrices*. Thesis (1932), University of Wisconsin; abstract in the Bulletin of the American Mathematical Society, vol. 38 (1932), p. 633.

¶ The dot sum representing A indicates that the blocks A_i occur along the principal diagonal of A and that all remaining blocks are zero matrices.

where $|A_i-\lambda I|=(a_i-\lambda)^{n_i}$, $a_i\neq a_j$ *in case* $i\neq j$, *and* $B=(B_{ij})$, *where* B_{ij} $(i, j=1, 2, \cdots, r)$ *are* $n_i\times n_j$ *matrices, have the property* P, *then the determinant*

$$(1) \qquad |B|=\prod_{i=1}^{r}|B_{ii}|$$

and the characteristic values of B *associated with* a_i *of* A *are those of* B_{ii}, *and moreover the matrices*

$$A_i^{\sigma}B_{ii}-B_{ii}A_i^{\sigma} \qquad (i=1, 2, \cdots, r; \sigma=0, 1, \cdots, n_i-1)$$

and

$$A_{\alpha}^{\sigma_\alpha}B_{\alpha\beta}A_{\beta}^{\sigma_\beta}B_{\beta\gamma}\cdots A_{\rho}^{\sigma_\rho}B_{\rho\alpha}A_{\alpha}^{\sigma_\alpha},$$

where $\alpha, \beta, \cdots, \rho$ *are not all equal and* $\sigma_k=0, 1, 2, \cdots, n_k-1$, *are nilpotent matrices.*

The proof of this theorem follows. Let

$$\psi_i(\lambda)=\prod_{j=1}^{i-1}(a_j-\lambda)^{n_j}\prod_{j=i+1}^{r}(a_j-\lambda)^{n_j} \qquad (i=1, 2, \cdots, r),$$

that is, $\psi_i(\lambda)$ is the characteristic function of A from which the factors corresponding to A_i are deleted. The matrix $\psi_i(A_i)$ is non-singular, for by hypothesis $a_i\neq a_j$, if $i\neq j$; on the other hand, $\psi_i(A_j)=0$, if $i\neq j$. Let $g_i(A_i)$ be the inverse of $\psi_i(A_i)$; then the polynomial $\phi_i(\lambda)\equiv\mu_i g_i(\lambda)\psi_i(\lambda)$ exists such that

$$\phi_i(A_i)=\mu_i I_i,$$
$$\phi_i(A_j)=0, \qquad i\neq j,$$

where μ_i is an arbitrary parameter, and I_i is a unit matrix of order n_i. Obviously $\phi_i(A)$ is the direct sum of zero matrices save the ith, which is $\mu_i I_i$. Then since A and B have the property P, the matrix $\phi_i(A)+B$ must have the characteristic value $\phi_i(a_i)+\beta=\mu_i+\beta$, where β is a characteristic value of B which is associated with a_i of A. Hence

$$|\phi_i(A)+B-(\mu_i+\beta)I|=0 \qquad (i=1, 2, \cdots, r),$$

for all values of μ_i. We shall display this determinant for the case $i=1$:

$$\Delta(\mu_1)=\begin{vmatrix} B_{11}-\beta I_1 & B_{12} & \cdots & B_{1r} \\ B_{21} & B_{22}-(\mu_1+\beta)I_2 & \cdots & B_{2r} \\ \cdot & \cdot & \cdots & \cdot \\ B_{r1} & B_{r2} & \cdots & B_{rr}-(\mu_1+\beta)I_r \end{vmatrix}.$$

The determinant is a polynomial of degree $n-n_1$ in μ_1, whose constant term, $\Delta(0)=|B-\beta I|$, vanishes because β is a characteristic value of B. Moreover, since $\Delta(\mu_1)$ is identically zero under the hypothesis that A and B have the property P, we conclude that

$$|B_{11} - \beta I_1| = 0,$$

since this is the coefficient of $\mu_1^{n-n_1}$. That is, $\Delta(\mu_1)$ vanishes identically only if all characteristic values of B_{11} are also characteristic values of B. Hence the n_1 characteristic values of B associated with a_1 are those of B_{11}. Similarly the characteristic values of B_{ii} $(i=2, 3, \cdots, r)$ are also characteristic values of B and are associated with A_i respectively. We can therefore conclude that equation (1) holds and the first part of the theorem above is proved.

To prove the final assertion of the theorem above, we build up further polynomials in A and B. Let

$$\begin{aligned} C^{(\sigma)} &= (\phi_1(A) + \phi_2(A) + \cdots + \phi_r(A))A^\sigma \\ &= \mu_1 A_1^\sigma + \mu_2 A_2^\sigma + \cdots + \mu_r A_r^\sigma; \end{aligned}$$

then

$$D^{(\sigma)} = C^{(\sigma)}B - BC^{(\sigma)} = (\mu_i A_i^\sigma B_{ij} - \mu_j B_{ij} A_j^\sigma)$$

is again a polynomial in A and B. Since A and B have the property P, the characteristic values of $D^{(\sigma)}$ are all zeros. Moreover, A and $D^{(\sigma)}$ have the property P; hence, according to the demonstration already given, the characteristic values of $D^{(\sigma)}$ associated with a_i of A are those of

$$D_{ii}^{(\sigma)} = \mu_i(A_i^\sigma B_{ii} - B_{ii}A_i^\sigma).$$

Since these are all zeros, the matrices

$$A_i^\sigma B_{ii} - B_{ii}A_i^\sigma \qquad (i = 1, 2, \cdots, r; \sigma = 0, 1, \cdots, n_i - 1)$$

must all be nilpotent. At most the first n_i powers of A_i are linearly independent, hence σ need not exceed $n_i - 1$.

Finally let

$$C_k^{(\sigma_k)} = \mu_{1k}A_1^{(\sigma_k)} + \mu_{2k}A_2^{(\sigma_k)} + \cdots + \mu_{rk}A_r^{(\sigma_k)},$$

$$D_k^{(0)} = C_k^{(0)}B - BC_k^{(0)},$$

where the μ_{ik} $(i=1, 2, \cdots, r; k=1, 2, \cdots, s)$ are arbitrary parameters. Then the matrix

$$C_1^{(\sigma_1)}BC_2^{(\sigma_2)}B \cdots C_{k-1}^{(\sigma_{k-1})}D_k^{(0)}C_{k+1}^{(\sigma_{k+1})} \cdots BC_s^{(\sigma_s)} \tag{2}$$

is again a polynomial in A and B and is nilpotent, because the characteristic values of $D_k^{(0)} = ((\mu_{ik} - \mu_{jk})B_{ij})$ are zeros. The characteristic values of the matrix (2) associated with a_α of A are those of an $n_\alpha \times n_\alpha$ matrix which is a linear combination of terms of the form

$$A_\alpha^{\sigma_1} B_{\alpha\beta} A_\beta^{\sigma_2} B_{\beta\gamma} \cdots A_\mu^{\sigma_{k-1}} B_{\mu\nu} A_\nu^{\sigma_{k+1}} \cdots B_{\rho\alpha} A_\alpha^{\sigma_s}$$

with independent coefficients in μ_{ij} $(i=1, 2, \cdots, r; j=1, 2, \cdots, s)$, and that where $\mu=\nu$ is identically zero since $D_{\nu\nu}^{(0)}=0$. Hence each of the terms above must be nilpotent, where not all $\alpha, \beta, \cdots, \rho$ are equal. This completes the proof of the theorem.

We may note here that the foregoing implies that all matrices

$$B_{\alpha\beta} B_{\beta\gamma} \cdots B_{\rho\alpha}$$

are nilpotent provided not all subscripts $\alpha, \beta, \gamma, \cdots, \rho$ are equal. If

$$A_\alpha^{\sigma_\alpha} B_{\alpha\alpha} - B_{\alpha\alpha} A_\alpha^{\sigma_\alpha} \text{ and } A_\alpha^{\sigma_\alpha} B_{\alpha\beta} A_\beta^{\sigma_\beta} B_{\beta\alpha}, \qquad \alpha \neq \beta,$$

are nilpotent for $(\alpha, \beta=1, 2, \cdots, r; \sigma_\alpha=0, 1, 2, \cdots, n_\alpha-1$ and $\sigma_\beta=0, 1, 2, \cdots, n_\beta-1)$, it is quite possible that A and B have the property P; though the writer has not succeeded in proving that such is the case nor that the conditions as stated in the theorem above are sufficient.

2. Sufficient conditions

We shall now develop sufficient conditions that the matrices A and B have the property P. To do so we shall find it convenient to make certain definitions of terms and to prove two lemmas.

Definition 2. *The matrix $M=(m_{ij})$ of order $p \times q$ has $p+q-1$ diagonals. These we number consecutively beginning with that containing the single element $m_{p,1}$. Then we say that the rth diagonal, $0 \leqq r \leqq p+q$, of M is starred if the first $r-1$ diagonals of M contain only zero elements and the remaining diagonals of M may or may not contain zero elements.*

That is to say, the diagonal to be starred of a $p \times q$ zero matrix may be chosen in $p+q+1$ ways; that of a non-zero $p \times p$ diagonal matrix may be chosen in p ways. For example, the 3×2 matrix may have its diagonal to be starred chosen in one of the following six ways:

$$\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ 0 & 0, \end{array} \quad \begin{array}{cc} 0 & * \\ 0 & 0 \\ 0 & 0, \end{array} \quad \begin{array}{cc} * & \cdot \\ 0 & * \\ 0 & 0, \end{array} \quad \begin{array}{cc} \cdot & \cdot \\ * & \cdot \\ 0 & *, \end{array} \quad \begin{array}{cc} \cdot & \cdot \\ \cdot & \cdot \\ * & \cdot, \end{array} \quad \begin{array}{cc} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot, \end{array}$$

where an asterisk represents an element in a starred diagonal and both asterisks and dots represent arbitrary elements which may be zeros. In the first we may regard the sixth diagonal as starred and in the last the zeroth diagonal as starred.

Definition 3. *The matrix* $X=(X_{ij})$, *where* X_{ij} $(i,j=1,2,\cdots,r)$ *are* $n_i\times n_j$ *matrices whose* ρ_{ij}*th* $(0\leqq\rho_{ij}\leqq n_i+n_j)$ *diagonals respectively are starred, is an umbral matrix, if*

$$\text{(a)}\qquad \rho_{ij}+\rho_{ji}=n_i+n_j \qquad (i,j=1,2,\cdots,r),$$

and if (b) *in any row of* X *exactly* s *elements are starred then exactly* $s-1$ *other rows of* X *have the corresponding* s *elements starred.*

For example, the matrix

$$U=\left(\begin{array}{ccc|cc|cc|c}
\alpha & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \beta \\
0 & \alpha & \cdot & \alpha & \cdot & \cdot & \cdot & 0 \\
0 & 0 & \alpha & 0 & \alpha & \beta & \cdot & 0 \\
\hline
0 & \alpha & \cdot & \alpha & \cdot & \cdot & \cdot & 0 \\
0 & 0 & \alpha & 0 & \alpha & \beta & \cdot & 0 \\
\hline
0 & 0 & \gamma & 0 & \gamma & \delta & \cdot & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \delta & 0 \\
\hline
\gamma & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \delta
\end{array}\right),$$

where the Greek letters represent the location of starred elements in starred diagonals and dots represent the location of arbitrary elements, is an umbral matrix. According to the definition above, the umbral matrix X may be chosen in

$$\frac{\prod_{i=1}^{r}(n_i+n_k+1)}{2n_k+1}$$

ways, where $n_k\geqq n_i$ $(i=1,2,\cdots,r)$; for evidently the diagonals to be starred can be chosen arbitrarily only in the matrices X_{ik} or in X_{ki} $(i=1,2,\cdots,r)$; the remaining blocks have their starred diagonals determined by conditions (a) and (b) of the definition above. Determinants of umbral matrices, in which all blocks are square and the principal diagonal of each is starred, were studied by Williamson.† The direct product $A<B>=(a_{ij}B)$, where the principal diagonal of B is starred, is an umbral matrix of the same type.

Definition 4. *Two umbral matrices* X *and* Y *are related if corresponding elements of* X *and* Y *are starred.*

---

† Williamson, *Latent roots of a matrix of a special type*, Bulletin of the American Mathematical Society, vol. 37 (1931), pp. 585–590.

LEMMA I. *If X is an umbral matrix then there exists an orthogonal matrix Q such that*

$$QXQ^T = ((\mathfrak{X}_{ij})), \tag{3}$$

where Q^T is the transpose of Q, and where the $\nu_i \times \nu_j$ matrices $\mathfrak{X}_{ij}$ $(i,j=1,2,\cdots,\rho)$ are zero if $i>j$, contain only starred elements of X if $i=j$, and contain only arbitrary non-starred elements of X if $i<j$.†

According to condition (a) of Definition 3 of an umbral matrix $X=(x_{ij})$ $(i, j=1, 2, \cdots, n)$, we conclude that all x_{ii} $(i=1, 2, \cdots, n)$ are starred; also that if x_{hk} is starred, then x_{kh} is likewise starred; and on the other hand if x_{hk} is not starred, then either x_{hk} or x_{kh} is a zero element. From these conclusions and by condition (b) of the definition of an umbral matrix, it follows that if s rows have s starred elements each, then the corresponding s columns also have s starred elements each, and the s^2 intersections of these s rows and s columns locate the only starred elements in these rows and columns. Moreover, these s rows have only zero elements in z columns and only non-starred elements in u columns (i.e., $z+u=n-s$); and the corresponding s columns have only zero elements in u rows and only non-starred arbitrary elements in z rows. That is, by an interchange of rows and the corresponding interchange of columns of X, the $s\times s$ minor determined by them may be brought into the principal diagonal of the transformed matrix with an $s\times z$ zero matrix on its left and a $u\times s$ zero matrix below it. The same may be done with all starred elements of X. Hence X may be transformed to a matrix having all starred elements and no other elements of X in non-overlapping square blocks along the principal diagonal and zero blocks below and to the left of the diagonal blocks.

The interchange of rows of a matrix can always be accomplished by multiplying it on the left by a matrix, Q, having only one non-zero element, namely unity, in each row and column. The corresponding interchange of columns results if the matrix be multiplied on the right by Q^T, the transpose of Q. Obviously $QQ^T=I$. Hence the lemma is proved.

The matrix Q, which transforms the umbral matrix, X, to the form established above plays an important role in the sequel and is not unique, if $r>1$. However, we may so choose Q that any s rows and corresponding s columns whose intersections determine an $s\times s$ block in the principal diagonal of QXQ^T are not interchanged among themselves. That is, the s^2 elements in question have the same relative positions with respect to each other in the

† Hereafter the double parentheses will indicate a matrix in which only the blocks $\mathfrak{X}_{ij}$, $i>j$, are necessarily zero matrices.

diagonal blocks of QXQ^T that they enjoy in X. This fact will be useful later.

On the basis of the above lemma we have at once the following

COROLLARY. *The determinant of an umbral matrix depends on and only on its starred elements.*

LEMMA II. *If $f(X, Y)$ is a polynomial in the related umbral matrices X and Y, then X, Y, and $f(X, Y)$ are related umbral matrices.*

The matrix Q, which transforms X to the form (3), transforms Y to the corresponding matrix

$$QYQ^T = ((\Upsilon_{ij})), \tag{4}$$

where Υ_{ij} are $\nu_i \times \nu_j$ matrices. But

$$Qf(X, Y)Q^T = f(QXQ^T, QYQ^T)$$

is a matrix, whose diagonal blocks are $f(\mathfrak{X}_{ii}, \Upsilon_{ii})$ $(i=1, 2, \cdots, \rho)$ and those below and to the left of them are zero blocks. Hence the transform $Q^T(\quad)Q$ transforms $Qf(X, Y)Q^T$ to $f(X, Y)$, which is therefore an umbral matrix related to X and to Y.

COROLLARY 1. *If X and Y are related umbral matrices which are transformed by the orthogonal matrix Q to the forms* (3) *and* (4) *respectively, then the determinant*

$$|f(X, Y)| = \prod_{i=1}^{\rho} |f(\mathfrak{X}_{ii}, \Upsilon_{ii})|$$

and consequently depends only on the starred elements of X and Y.

COROLLARY 2. *If X is an umbral matrix and if all starred elements in and above or in and below its principal diagonal are zeros, then X is a nilpotent matrix.*

If all starred elements of X in and below its principal diagonal are zeros, then the orthogonal matrix Q can be so chosen that QXQ^T has only zero elements in and below its principal diagonal, and is therefore nilpotent. If all starred elements of X in and above the principal diagonal are zeros, the same conclusion holds.

COROLLARY 3. *If X and Y are two matrices of order n, if R exists such that RXR^{-1} and RYR^{-1} are related umbral matrices and RXR^{-1} is in the Jordan canonical form, and if all starred elements of RYR^{-1} in and above or in and below its principal diagonal are zeros, then $|X+Y| = |X|$.*

In the present case, where RXR^{-1} is in the Jordan canonical form, Q exists such that the diagonal blocks of $QRXR^{-1}Q^T$ are diagonal matrices having non-zero elements only in their principal diagonals since the only non-zero starred elements are those in the principal diagonal of RXR^{-1}. The diagonal blocks of $QRYR^{-1}Q^T$, under the hypotheses of the above corollary, have only zero elements in and below (or above) their principal diagonals, consequently $|X+Y| = |X|$. This corollary is a generalization of Frobenius' theorem† which requires X and Y to be commutative.

THEOREM II. *If A and B are given matrices of order n, if $RAR^{-1}=\bar{A}$ and $RBR^{-1}=\bar{B}$, where*

$$\bar{A} = A_1 \dotplus A_2 \dotplus \cdots \dotplus A_r,$$

and

$$A_i = \begin{pmatrix} a_i & 1 & 0 \cdots 0 \\ 0 & a_i & 1 \cdots 0 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ 0 & 0 & 0 \cdots a_i \end{pmatrix} \qquad (i = 1, 2, \cdots, r),$$

are matrices of order n_i, and if $\bar{A}$ and $\bar{B}$ are related umbral matrices, such that for every starred element b_{hk}^* of $B=(b_{ij})$ we have

$$b_{hk}^*(a_{hh} - a_{kk}) = 0, \tag{5}$$

either for all $h>k$, or for all $h<k$, where we designate the hth principal diagonal element of $\bar{A}$ by a_{hh} $(h=1, 2, \cdots, n)$, then A and B have the property P.

This theorem says merely that if for example

$$\bar{A} = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix} \dotplus \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} \dotplus \begin{pmatrix} b & 1 \\ 0 & b \end{pmatrix} \dotplus (b)$$

and U is the matrix given above (page 238), then $\bar{A}$ and U are related umbral matrices and have the property P in case $a=b$; if, however, $a\neq b$ and if either the β's or the γ's of U represent zero elements, then $\bar{A}$ and U have the property P.

Let Q be the matrix which transforms the related umbral matrices $\bar{A}$ and $\bar{B}$ to the forms

$$\begin{aligned} Q\bar{A}Q^T &= ((\mathfrak{A}_{ij})), \\ Q\bar{B}Q^T &= ((\mathfrak{B}_{ij})), \end{aligned} \tag{6}$$

where the diagonal blocks $\mathfrak{A}_{ii}$ and $\mathfrak{B}_{ii}$ $(i=1, 2, \cdots, \rho)$ are of order $\nu_i \times \nu_i$, $\sum_{i=1}^{\rho} \nu_i = n$. The matrices $\mathfrak{A}_{ii}$ are diagonal matrices with certain permutations

† Frobenius, *Über lineare Substitutionen und bilineare Formen*, Crelle's Journal, vol. 84 (1878), pp. 1–63.

of ν_i elements a_j $(j=1, 2, \cdots, r)$ in their principal diagonals. If all $a_j=a$ $(j=1, 2, \cdots, r)$, then all blocks $\mathfrak{A}_{ii}(i=1, 2, \cdots, \rho)$ are scalar matrices and as such have the property P with the matrices $\mathfrak{B}_{ii}$ $(i=1, 2, \cdots, \rho)$ respectively. Hence according to Lemmas I and II, the matrices A and B themselves have the property P without further restrictions upon B. However, if the a_j $(j=1, 2, \cdots, r)$ are not all equal, then neither the sum nor the product of $\mathfrak{A}_{ii}$ and $\mathfrak{B}_{ii}$ will necessarily have the property P; hence in order that the matrices A and B may have the property P further restrictions must be placed upon B.

As was pointed out following the proof of Lemma I, we may so choose Q that starred elements of $\overline{B}$ do not change the relative positions in the blocks $\mathfrak{B}_{ii}$ that they enjoy in $\overline{B}$; moreover, it is no restriction upon A to assume that in $\overline{A}$ all blocks A_j $(j=1, 2, \cdots, r)$ having the same characteristic value are adjacent. Hence with Q so chosen, like characteristic values a_j will occur in adjacent positions along the principal diagonals of each block $\mathfrak{A}_{ii}$ $(i=1, 2, \cdots, \rho)$. Now according to hypothesis b_{hk}^* $(a_{hh}-a_{kk})=0$, we have $b_{hk}^*=0$, if $a_{hh}\neq a_{kk}$, and the same is true of the corresponding blocks $\mathfrak{A}_{ii}$ and $\mathfrak{B}_{ii}$. For example, $\mathfrak{A}_{ii}$ and $\mathfrak{B}_{ii}$ under the hypothesis above must have the following forms:

$$\mathfrak{A}_{ii} = \begin{pmatrix} a_\alpha & 0 & 0 & 0 & 0 \\ 0 & a_\alpha & 0 & 0 & 0 \\ 0 & 0 & a_\beta & 0 & 0 \\ 0 & 0 & 0 & a_\beta & 0 \\ 0 & 0 & 0 & 0 & a_\gamma \end{pmatrix}, \qquad \mathfrak{B}_{ii} = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot \end{pmatrix},$$

where a_α, a_β, and a_γ are distinct and the dots of $\mathfrak{B}_{ii}$ represent arbitrary elements. Obviously matrices of this form have the property P. Hence if $b_{kh}^*(a_{hh}-a_{kk})=0$ for all $h>k$ the matrices $\overline{A}$ and $\overline{B}$ and consequently A and B have the property P. The same conclusion holds if $b_{hk}^*(a_{hh}-a_{kk})=0$ for all $h<k$. Herewith the theorem is proved.

THEOREM III. *If $Y=(Y_{ij})$, where Y_{ij} $(i, j=1, 2, \cdots, r)$ are $n_i\times n_j$ matrices having zero elements in the first $[n_i, n_j]-1$ diagonals where $[n_i, n_j]$ is the greater of the two numbers n_i and n_j or their common value, then the determinant of Y is dependent only on the diagonal elements of those blocks Y_{ij}, where $n_i=n_j$; that is,*

$$|Y| = |Y_\alpha| \cdot |Y_\beta| \cdot \quad \cdots \quad \cdot |Y_\lambda|,$$

where $Y_k=(Y_{ij})$ $(k=\alpha, \beta, \cdots, \lambda)$, and i, j run only over those values for which Y_{ij} are square matrices of order n_k, and $n_\alpha, n_\beta, \cdots, n_\lambda$ are the distinct values of $n_i(i=1, 2, \cdots, r)$.

It is no restriction upon Y to assume that $n_1 \geqq n_2 \geqq \cdots \geqq n_r$. Now if we regard the n_ith diagonal of the blocks Y_{ij} as starred, then Y is an umbral matrix, and by Lemma I and its corollary the determinant of Y depends only on the n_ith diagonal elements of its blocks. However, the elements of the n_ith diagonal of Y_{ij} are all zeros in case $i>j$ and $n_i<n_j$. Hence by a procedure similar to that followed in the proof of Theorem II, we can show that the present theorem holds.†

The matrices commutative or quasi-commutative with the Jordan canonical matrix A given in Theorem II are of the form Y here considered save that the blocks Y_{ij} must be zero if $a_i \neq a_j$ and the non-zero elements of a given diagonal of Y_{ij} are not linearly independent in such cases. The above theorem makes no such restrictions upon the elements in and above the $[n_i, n_j]$th diagonals of Y_{ij}.

Added in proof, January 7, 1936. Another attack upon the problem here discussed was recently made by McCoy (Bulletin of the American Mathematical Society, vol. 41 (1935), p. 635, abstract 41-9-351). Professor McCoy has kindly communicated a more complete statement of his main results to me. They are very general and supplement rather than duplicate those we give above.

---

† Williamson, Bulletin of the American Mathematical Society, vol. 37 (1931), pp. 585–590, proved a special case under the above theorem.

University of Wisconsin, Extension Division,
Milwaukee, Wis.