SOME THEOREMS ON MATRICES WITH REAL QUATERNION ELEMENTS

N. A. WIEGMANN

1. Introduction. Matrices with real quaternion elements have been dealt with in earlier papers by Wolf **(10)** and Lee **(4)**. In the former, an elementary divisor theory was developed for such matrices by using an isomorphism between $n \times n$ real quaternion matrices and $2n \times 2n$ matrices with complex elements. In the latter, further results were obtained (including, mainly, the transforming of a quaternion matrix into a triangular form under a unitary similarity transformation) by using a different isomorphism. Certain other related results have also been obtained **(1)**. Others, including Moore and Ingraham, have considered quaternion matrices earlier.

The intent here is to consider how other theorems which hold for matrices in the complex field may hold for quaternion matrices. To do this, the isomorphism in **(4)** is employed. First, an analog of the Jordan normal form is obtained; this result is closely related, of course, to the final result in **(10)** concerned with necessary and sufficient conditions for similarity of quaternion matrices, but here a proof is employed which depends entirely on known complex matrix theory, which throws light on the structure of the similarity transformation, and which leads in a natural way to a definition of elementary divisors for quaternion matrices. Next, this Jordan form is used to obtain some results concerning commutative matrices. In part 4, the familiar polar form of a complex matrix is shown to hold in the quaternion case. Next, some further properties of normal quaternion matrices are verified and, in the final section, some properties of quaternion matrices relative to unitary (quaternion) equivalence transformations are obtained.

2. An analog of the Jordan normal form. Let the $n \times n$ quaternion matrix A be written in the form $A = A_1 + jA_2$ where A_1 and A_2 are (uniquely determined) matrices with complex elements (where every quaternion element is considered as written in the form $a = (a_1 + a_2 i) + j(a_3 + a_4 i)$ where each a_i is real). Form the $2n \times 2n$ complex matrix

$$A^* = \begin{bmatrix} A_1 & -A_2{}^C \\ A_2 & A_1{}^C \end{bmatrix}$$

(where A^C denotes the matrix obtained by taking the complex conjugate of each element of A and, later, A^{CT} denotes the transpose of A^C). According to **(4)**, the correspondence between A and A^* is an isomorphism and has properties as developed there.

Received April 30, 1954.

If it is possible to show that for a given A^* there is determined an $n \times n$ matrix J_1 in the (complex) Jordan normal form such that a non-singular matrix P and a matrix J exist so that $A^*P = PJ$ where P and J have the forms, respectively,

$$\begin{bmatrix} P_1 & -P_2^C \\ P_2 & P_1^C \end{bmatrix} \text{ and } \begin{bmatrix} J_1 & 0 \\ 0 & J_1^C \end{bmatrix},$$

then an analog of the Jordan normal form can be obtained for the quaternion matrix A. For it can be easily seen that the inverse of P must also be of the form

$$\begin{bmatrix} Q_1 & -Q_2^C \\ Q_2 & Q_1^C \end{bmatrix},$$

so that $P^{-1}A^*P = J$. But according to the nature of the correspondence between A^* and A this implies that $(Q_1+jQ_2)(A_1+jA_2)(P_1+jP_2) = J_1+j\cdot 0$ or that $(Q_1 + jQ_2)A(P_1 + jP_2) = J_1$ where $(Q_1 + jQ_2)(P_1 + jP_2) = I$, so that A is similar to a complex $n \times n$ Jordan form J_1.

Since A^* is a matrix with complex elements, there exists a non-singular matrix P such that $P^{-1}A^*P = J$ is the Jordan form of A^* so that $A^*P = PJ$. In the following steps it will be shown that a P and a J can be obtained which satisfy this relation and are of the desired form.

(a) Let $\alpha_1, \ldots, \alpha_m$ be the distinct (complex) characteristic roots of A^*. Then each column of P is a column vector v with $2n$ elements satisfying one and only one of the following relations:

$$\text{(i) } A^*v = v\alpha_i,$$
$$\text{(ii) } A^*v = w + v\alpha_i,$$

where w is the column vector adjacent to v on the left. All $2n$ column vectors are linearly independent and for each α_i there exists at least one column of the type (i).

(b) For simplicity in notation let α be a root for which $v_1, v_2, \ldots, v_t$ are the set of column vectors of P of type (i) relative to α. Let the column vector v_1^* be defined relative to v_1 as follows: If v_1 is a column vector whose transpose is the row vector

$$[v_{11}, v_{21}, \ldots, v_{n1}, w_{11}, w_{21}, \ldots w_{n1}],$$

then v_1^* is the column vector whose transpose is the row vector

$$[-\bar{w}_{11}, -\bar{w}_{21}, \ldots -\bar{w}_{n1}, \bar{v}_{11}, \bar{v}_{21}, \ldots \bar{v}_{n1}].$$

If v_1 is not the zero vector, then v_1 and v_1^* are linearly independent, for if $c_1v_1 + c_2v_1^* = 0$, it follows that

$$(c_1\bar{c}_1 + c_2\bar{c}_2)w_{k1} = 0, \quad (c_1\bar{c}_1 + c_2\bar{c}_2)v_{k1} = 0 \quad (k = 1, 2, \ldots, n),$$

so that $c_1 = c_2 = 0$. Also, if $A^*v_1 = v_1\alpha$, it follows that $A^*v_1^* = v_1^*\bar{\alpha}$. Let us consider, first, vectors of type (i).

https://doi.org/10.4153/CJM-1955-024-x Published online by Cambridge University Press

Let α be real. Then v_1 and v_1^* are linearly independent vectors of type (i) for α, and either exhaust the number of such linearly independent column vectors or there exists another, say v_2, which is linearly independent of v_1 and v_1^*. Form v_2^*; then $v_1, v_1^*, v_2,$ and v_2^* are linearly independent, for if $c_1v_1 + c_2v_1^* + c_3v_2 + c_4v_2^* = 0$, then $\bar{c}_1v_1^* - \bar{c}_2v_1 + \bar{c}_3v_2^* - \bar{c}_4v_2 = 0$. But by properly combining these relations, it would follow that

$$(\bar{c}_3c_1 + c_4\bar{c}_2)v_1 + (\bar{c}_3c_2 - c_4\bar{c}_1)v_1^* + (c_3\bar{c}_3 + c_4\bar{c}_4)v_2 = 0,$$

so that $c_3 = c_4 = 0$ and so, from above, $c_1 = c_2 = 0$. Either v_1, v_1^*, v_2, v_2^* exhaust the number of linearly independent vectors of type (i) for α, or they do not. By means of this process there is obtained a set of linearly independent vectors of the form $v_1, v_1^*, \ldots, v_k, v_k^*$ which provide a basis for the vectors of type (i) corresponding to each real α.

Let α be non-real complex. Then if the matrix P contains a set of vectors $v_1, v_2, \ldots, v_t$ such that $A^*v_j = v_j\alpha$ $(j = 1, 2, \ldots, t)$, it follows that $A^*v_j^* = v_j^*\bar{\alpha}$ $(j = 1, 2, \ldots, t)$ (where the v_j^* are linearly independent since the v_j are), that there are no other vectors linearly independent of these for which this is true, and that $\bar{\alpha}$ is also a root of A^*.

Since $\alpha_1, \alpha_2, \ldots, \alpha_m$ are distinct, the sets of linearly independent vectors of type (i) obtained in this way are linearly independent and are equal in number to those column vectors of type (i) in the matrix P.

(c) Consider vectors v of type (ii); these may be written as $(A^* - \alpha_iI)v = w$; and it follows that $(A^* - \bar{\alpha}_iI)v^* = w^*$.

Let α be real. Let there be $2k$ vectors $v_1, v_2, \ldots, v_p, \ldots, v_{2k}$ of P of type (i) corresponding to α and let them be written in such an order that if there exist for some v_i vectors $v_i^{(1)}$ of P so that

(iii) $$(A^* - \alpha I)v_i^{(1)} = v_i,$$

these vectors, $v_1, v_2, \ldots, v_p$ are written together and first in this ordering. Then p must be even and a set of linearly independent vectors of the above v, v^* type can be obtained which span the same space as $v_1, v_2, \ldots, v_p$. For if $(A^* - \alpha I)v_1^{(1)} = v_1$, then $(A^* - \alpha I)v_1^{(1)*} = v_1^*$, and since v_1 and v_1^* are linearly independent, $v_1^{(1)}$ and $v_1^{(1)*}$ are also. Either $p = 2$, or the process can be continued as before, so that $p = 2q$. In this way we see that there exists a set of linearly independent vectors, $v_1, v_1^*, \ldots, v_q, v_q^*, \ldots, v_k, v_k^*$ (which form a basis for all vectors of type (i) corresponding to α) such that $v_1^{(1)}$, $v_1^{(1)*}, \ldots, v_q^{(1)}$ $v_q^{(1)*}$, provide a basis for the space spanned by $v_1^{(1)}, \ldots, v_{2q}^{(1)}$ as taken above where v_i and $v_i^{(1)}$ are related as above. If for some of the $v_j^{(1)}$ there exist $v_j^{(2)}$ in P such that

$$(A^* - \alpha I)v_j^{(2)} = v_j^{(1)},$$

the above process can be repeated, and a set of vectors, taken notationally as $v_1^{(2)}, v_1^{(2)*}, \ldots, v_s^{(2)}, v_s^{(2)*}$ (which span the same space as the linearly independent $v_j^{(2)}$), is obtained which stand in the same relation to $v_1^{(1)}, v_1^{(1)*}, \ldots, v_q^{(1)}$, $v_q^{(1)*}$ as the latter do to $v_1, v_1^*, \ldots, v_k, v_k^*$.

If α is non-real complex, let $v_1, v_2, \ldots, v_p, \ldots, v_t$ (the vectors of type (i) corresponding to α) be ordered in such a way that for $v_1, v_2, \ldots, v_p$ there exist $v_1^{(1)}, v_2^{(1)}, \ldots, v_p^{(1)}$ satisfying (iii). Then $(A^* - \bar{\alpha}I)v_i^{(1)*} = v_i^*$, $i = 1,2, \ldots, p$, and $v_1^{(1)*}, v_2^{(1)*}, \ldots, v_p^{(1)*}$ are such that there exists no vector, w, linearly independent of them such that $(A^* - \bar{\alpha}I)w$ is in the space generated by $v_1^*, v_2^*, \ldots, v_p^*$. For some $v_i^{(1)}$ there may exist $v_i^{(2)}$ such that

$$(A^* - \alpha I)v_i^{(2)} = v_i^{(1)};$$

in this way a set of $v_i^{(2)*}$ are determined and the process is seen to be a general one.

(d) By the above, a set of $2n$ linearly independent vectors, taken notationally as $w_1, w_2, \ldots, w_n, w_1^*, w_2^*, \ldots, w_n^*$, are obtained such that either $A^*w_i = w_i\alpha$ (and so $A^*w_i^* = w_i^*\bar{\alpha}$, for any α), or, for some w_i satisfying $A^*w_i = w_i\alpha$, there exist among the above $2n$ vectors certain vectors taken notationally as $w_i^{(1)}, w_i^{(2)}, \ldots, w_i^{(s)}$ such that

$$\begin{aligned} (A^* - \alpha I)w_i^{(1)} &= w_i, \\ (A^* - \alpha I)w_i^{(j)} &= w_i^{(j-1)}, \qquad j = 2, 3, \ldots, s; \end{aligned}$$

in this case it follows that

$$\begin{aligned} (A^* - \bar{\alpha}I)w_i^{(1)*} &= w_i^*, \\ (A^* - \bar{\alpha}I)w_i^{(j)*} &= w_i^{(j-1)*}, \qquad j = 2, 3, \ldots, s. \end{aligned}$$

It is now evident that by properly arranging the w_i and w_i^*, a $2n \times 2n$ matrix P can be obtained such that $A^*P = PJ$ as indicated above. If in J_1 (as used there), the roots $\alpha = a + bi$ are such that $b \geqq 0$, then a canonical form has been obtained for A^* and hence for the quaternion matrix A.

THEOREM 1. *Every $n \times n$ matrix with real quaternion elements is similar under a matrix transformation with real quaternion elements to a matrix in (complex) Jordan normal form with diagonal elements of the form $a + bi$, $b \geqq 0$.*

3. Properties of commutative matrices. According to a theorem due to Taber **(5)**, if a matrix A with complex elements is non-derogatory, the only matrices commutative with A are polynomial functions of A. An equivalent theorem had been previously given by Frobenius **(3**, Theorem XIII**)**.

In order to obtain an analog for this theorem where A contains real quaternion elements, let such a matrix A be defined to be non-derogatory when its Jordan normal form matrix (as obtained in the preceding) is non-derogatory.

Let A and B be quaternion matrices such that $AB = BA$ where A is non-derogatory. Let $PAP^{-1} = J = J_1 \dotplus J_2 \dotplus \ldots \dotplus J_m$ be the Jordan form of A where:

$$J_i = \begin{bmatrix} \alpha_i & 1 & 0 & \ldots & 0 \\ 0 & \alpha_i & 1 & \ldots & 0 \\ & \ldots & & \ldots & \\ 0 & \ldots & & \alpha_i & 1 \\ 0 & \ldots & & 0 & \alpha_i \end{bmatrix}, \qquad i = 1,2, \ldots, m,$$

where $\alpha_i \neq \alpha_j$ when $i \neq j$ and $\alpha_k = a_k + ib_k$, $b_k \geqq 0$. Let $PBP^{-1} = B_1$ so that $JB_1 = B_1J$.

LEMMA. *$B_1 = B_{11} \dotplus B_{12} \dotplus \ldots \dotplus B_{1m}$ where B_{1i} has the same order as J_i, where*

$$B_{1i} = \begin{bmatrix} b_{i1} & b_{i2} & b_{i3} & & \cdots & & b_{in} \\ 0 & b_{i1} & b_{i2} & & \cdots & & \\ 0 & 0 & b_{i1} & & \cdots & & \\ & \cdots & & & \cdots & & \\ & \cdots & & b_{i1} & & b_{i2} & b_{i3} \\ & \cdots & & 0 & & b_{i1} & b_{i2} \\ 0 & \cdots & & 0 & & 0 & b_{i1} \end{bmatrix}$$

and where

(i) *if α_i is real, the non-diagonal elements of B_{1i} are quaternions while the diagonal elements are complex;*

(ii) *if α_i is non-real complex, the elements of the corresponding B_{1i} are complex.*

The following may be noted: if α is a non-real complex element, b a quaternion element, and $\alpha b = b\alpha$, then b is a complex number; if α is non-real complex, if β is complex, and c a quaternion element such that $\alpha c = c\alpha + \beta$, then c is complex and $\beta = 0$. Also no α_i, above, is the conjugate of an α_j.

Let $B_1 = (b_{ij})$, let J_1 be of order $r \times r$, and consider the upper left $r \times r$ sub-matrix of the product $JB_1 = B_1J$. The following relations result:

$$\begin{aligned} &\alpha_1 b_{r1} = b_{r1}\alpha_1, && \\ &a_1 b_{ri} = b_{r\,i-1} + b_{ri}\alpha_1, && i = 2, 3, \ldots, r, \\ &\alpha_1 b_{i1} + b_{i+1,1} = b_{i1}\alpha_1, && i = 1, 2, \ldots, r-1, \\ &\alpha_1 b_{it} + b_{i+1,t} = b_{i,t-1} + b_{it}\alpha_1, && \begin{cases} t = 2, 3, \ldots, r, \\ i = 1, 2, \ldots, r-1. \end{cases} \end{aligned}$$

If α_1 is real then, although the b_{ij} are quaternion elements, all commutative properties hold for these relations (as in the complex case as treated by Taber) and the upper left $r \times r$ matrix has the form B_{11} with all quaternion elements, in general. If α_1 is non-real complex, it follows from the first relation that b_{r1} is complex; from this and the third relation it follows that all elements in the first column above b_{r1} are complex (and in fact, except for b_{11}, all are 0); from the second relation it can be seen that all elements of the rth row of this submatrix are 0 except b_{rr} which is complex. Using the fourth set of relations, we see that the remaining elements are complex, all necessary commutative properties hold, and that the submatrix has the B_{11} form. B_{11} now has the required form unless, for a real α_1, the diagonal elements are quaternions; if so, there exists a quaternion element b such that $bb_{i1}\bar{b} = \beta$ is a complex number where $b\bar{b} = 1$. Form the $n \times n$ matrix $Q = bI_1 \dotplus I_2$ where I_1 and I_2 are identity matrices and I_1 is of order $r \times r$. Then $Q^{-1} = \bar{b}I_1 \dotplus I_2$ and QB_1Q^{-1} has the form required and $QJQ^{-1} = J$.

Let J_2 be of order $s \times s$, and consider the $s \times r$ submatrix directly below B_{11} in the matrix B_1. Upon comparing corresponding elements of this $s \times r$ submatrix in the product $JB_1 = B_1J$, we see that the set of following relations appear:

$$\begin{aligned}
&\alpha_2 b_{r+s,1} = b_{r+s,1}\alpha_1, && \\
&\alpha_2 b_{r+s,i} = b_{r+s,i-1} + b_{r+s,i}\alpha_1, && i = 2, \ldots, r, \\
&\alpha_2 b_{i1} + b_{i+1,1} = b_{i1}\alpha_1, && i = r+1, \ldots, r+s-1, \\
&\alpha_2 b_{it} + b_{i+1,t} = b_{i,t-1} + b_{it}\alpha_1, && \begin{cases} t = 2, 3, \ldots, r, \\ i = r+1, r+2, \ldots r+s-1. \end{cases}
\end{aligned}$$

Since, for $i \neq j$, $\alpha_i \neq \alpha_j$ and $\alpha_i \neq \bar{\alpha}_j$, it follows from these relations that all elements of this $s \times r$ submatrix of B_1 are zero. In this way it can be shown that $B_1 = B_{11} \dot{+} B_2$ where B_{11} has the form given in the lemma. When B_2 is treated in like fashion, the lemma follows.

Consider next the possibility of representing this B_1 as a polynomial in J_1 where J_1 contains only complex elements. It is evident (from the work of Taber or by merely considering the set of equations obtained) that it is possible to determine two sets, x_i and x_i', $i = 0, 1, 2, \ldots, n-1$, of quaternion elements such that

$$B_1 = \sum_{i=0}^{n-1} x_i J^i = \sum_{i=0}^{n-1} J^i x_i'.$$

If all the diagonal elements of J are real, $x_i J^i = J^i x_i$; if all the diagonal elements of J are non-real complex, all elements of B_1 are complex and so are the x_i so that again $x_i J^i = J^i x_i$; and the same would be true if all the elements of B_1 were complex regardless of the nature of the α_i in J. In these instances if $x_j = \rho_j u_j$ (where ρ_j is the real absolute value of the quaternion element x_j and u_j the related quaternion of absolute value one), then

$$B = P^{-1}B_1P = \sum_{i=0}^{n-1} \rho_i P^{-1}(u_i I)P \cdot P^{-1}(J^i)P = \sum_{i=0}^{n-1} \rho_i U_i A^i$$

where $U_i = P^{-1}(u_i I)P$ and $U_i A = A U_i$ for each i. It follows that:

THEOREM 2. *If A and B are quaternion matrices, if $AB = BA$, and if A is non-derogatory with either all real or all non-real complex roots, then*

$$B = \sum_{i=0}^{n-1} \rho_i U_i A^i$$

where the ρ_i are real, $U_i A = A U_i$ for each i, and each U_i has a single characteristic root of absolute value one.

4. A polar form. Every complex number has the familiar polar form $\rho e^{i\theta}$ and, as has been seen, the same is true for a quaternion. For a matrix A with complex elements a polar representation has been obtained when A is non-singular by Wintner and Murnaghan **(9)** and when A is singular by Williamson **(7)**. It also exists for quaternion matrices according to the following:

THEOREM 3. *Every $n \times n$ matrix A with real quaternion elements can be expressed as $A = H_1 W_1 = W_1 K_1$ where H_1 and K_1 are hermitian (quaternion) matrices and W_1 is a unitary (quaternion) matrix; if A is non-singular the representation is unique, and if A is singular, H_1 and K_1 are unique but W_1 is arbitrary to some extent.*

Let $A = A_1 + jA_2$ where A_1 and A_2 are (as in §2) uniquely determined matrices with complex elements. Then A is isomorphic to A^* where

$$\text{(i)} \qquad A^* = \begin{bmatrix} A_1 & -A_2{}^C \\ A_2 & A_1{}^C \end{bmatrix}.$$

Since A^* has complex elements $A^* = HU = UK$ by **(9)** and **(7)**, where H and K are hermitian and U unitary. Then AA^{CT} corresponds to H^2 and there exists a unitary quaternion matrix (see **(4)**, for example) $V_3 = V_1 + jV_2$ so that $V_3 AA^{CT} V_3{}^{CT} = D$ is a diagonal matrix with real elements and, consequently, if

$$V = \begin{bmatrix} V_1 & -V_2{}^C \\ V_2 & V_1{}^C \end{bmatrix},$$

then $VA^*A^{*CT}V^{CT} = D \dotplus D$. Since H is hermitian with non-negative real roots, there exists a unitary matrix W such that $WHW^{CT} = D_1$ is diagonal with these non-negative real roots along the diagonal; and this W can be chosen in such a way that

$$WH^2W^{CT} = WA^*A^{*CT}W^{CT} = D \dotplus D$$

so that $D_1{}^2 = D \dotplus D$ and so $D_1 = D_2 \dotplus D_2$ where the diagonal elements of D_2 are the positive square roots of the corresponding real roots of D. Then H must be of the same form as A^* in (i) for if $X = VW^{CT}$, then

$$XWA^*A^{*CT}W^{CT}X^{CT} = X(D \dotplus D)X^{CT} = VA^*A^{*CT}V^{CT} = D \dotplus D$$

so that $X(D \dotplus D) = (D \dotplus D)X$ and so $X(D_2 \dotplus D_2) = (D_2 \dotplus D_2)X$. From this, $XWHW^{CT}X^{CT} = XD_1X^{CT} = D_1 = VHV^{CT}$ so that $H = V^{CT}D_1V$ and from the form of the matrices on the right side of this equality, their product is of type (i).

From $A^* = HU$, it follows that $VA^*V^{CT} = VHV^{CT}VUV^{CT}$ where the matrices have the form

$$\begin{bmatrix} B_1 & -B_2{}^C \\ B_2 & B_1{}^C \end{bmatrix} = \begin{bmatrix} D_2 & 0 \\ 0 & D_2 \end{bmatrix} \begin{bmatrix} U_1 & U_2 \\ U_3 & U_4 \end{bmatrix}.$$

If A^* is non-singular, D_2 is non-singular and by equating corresponding block matrices, $U_4 = U_1{}^C$ and $U_2 = -U_3{}^C$.

If A^* is singular (in which case there is some arbitrariness involved in the choice of U in HU), then D_2 is singular; let the first r diagonal elements be non-zero, the remaining being 0. From this it can be seen that $D_2(U_1 - U_4{}^C) = 0$ and $D_2(U_3 + U_2{}^C) = 0$; this means that the first r rows of U_4 are the conjugates of the first r rows of U_1 and the first r rows of U_2 are the negative conjugates of the first r rows of U_3. Since VUV^{CT} is unitary, these $2r$ rows are linearly independent and by means of the v, v^*-basis procedure employed in §2 above, it is seen that it is possible to complete the remaining rows of this matrix so that it is unitary and of the form (i). From the form of each matrix in the $2n \times 2n$ matrix relation $A^* = HU$, it follows that A can be expressed as required by the theorem. Since $U^{CT}A^* = U^{CT}HU = K$ is hermitian, $A^* = UK$ holds (uniquely if A^* is non-singular) and the theorem is true.

5. Properties of normal quaternion matrices. If A is a normal quaternion matrix, it can be brought into diagonal form under a unitary similarity transformation (see **(4)**, for example). Some further properties of normal quaternion matrices are verified here.

It is known that a complex matrix A is normal if and only if A^{CT} is a polynomial in A. If A is a normal quaternion matrix, there exists a unitary quaternion matrix U such that $UAU^{CT} = D$ where the characteristic roots of A appear in the diagonal matrix D. If $\alpha_1, \alpha_2, \ldots, \alpha_m$ are the distinct roots of A, the set of equations

$$\bar{\alpha}_i = \sum_{j=0}^{m-1} x_j \alpha_i^{\,j}, \qquad i = 1, 2, \ldots, m,$$

in x_j always have solutions in the complex field. This implies that

$$D^{CT} = \sum_{j=0}^{m-1} x_j D^j$$

and, if $x_j = \rho_j \cdot e^{i\theta_j}$,

$$A^{CT} = \sum_{j=0}^{m-1} \rho_j U(e^{i\theta_j} I)\, U^{CT} A^j = \sum_{j=0}^{m-1} \rho_j V_j A^j$$

where $V_j = U(e^{i\theta_j}I)U^{CT}$ is unitary and $V_jA = AV_j$ for all j. If more latitude is allowed for the degree of the polynomial, let the distinct roots be written in the form $\alpha_1, \alpha_2, \ldots, \alpha_r, \ldots, \alpha_m$ where $\alpha_1, \ldots, \alpha_r$ are the non-real complex roots. Let the roots $\alpha_1, \alpha_2, \ldots, \alpha_r, \ldots \alpha_m, \bar{\alpha}_1, \ldots, \bar{\alpha}_r$ be used to form the $m + r$ equations

$$\bar{\beta}_i = \sum_{j=0}^{m+r-1} x_j \beta_i^{\,j}$$

where β_i runs through the latter set of α_i and $\bar{\alpha}_i$; in this case the x_j will all be real and it follows that:

THEOREM 4. *A quaternion matrix A is normal if and only if A^{CT} is a polynomial in A with real coefficients.*

The following theorem will now be shown to hold as in the complex case:

THEOREM 5. *Two normal quaternion matrices A and B are commutative if and only if they can be diagonalized by the same unitary transformation.*

If $AB = BA$, let $UAU^{CT} = D$ where D is diagonal such that like roots are in consecutive order, with real roots $\alpha_1, \ldots, \alpha_s$ first, and complex roots $\beta_1, \ldots, \beta_t$, $(\beta_k = \gamma_k + i\delta_k,\ \delta_k > 0)$ next. Let

$$\begin{bmatrix} D & 0 \\ 0 & D^C \end{bmatrix} \text{ and } \begin{bmatrix} C_1 & -C_2{}^C \\ C_2 & C_1{}^C \end{bmatrix}$$

be the $2n \times 2n$ complex matrices which are isomorphic to D and UBU^{CT}, respectively. From the commutative property $DC_1 = C_1D$ and $D^CC_2 = C_2D$, and so

$$\begin{aligned} C_1 &= C_{11} \dotplus \ldots \dotplus C_{1s} \dotplus C'_{11} \dotplus \ldots \dotplus C'_{1t}, \\ C_2 &= C_{21} \dotplus \ldots \dotplus C_{2s} \dotplus 0 \quad \dotplus \ldots \dotplus 0, \end{aligned}$$

where $D = \alpha_1 I_1 \dotplus \ldots \dotplus \alpha_s I_s \dotplus \beta_1 I'_1 \dotplus \ldots \dotplus \beta_t I'_t$ and where C_{1j} and C_{2j} have the same order as the identity matrix I_j and C'_{1j} and the corresponding 0 matrix in C_2 have the same order as the identity matrix I'_j. Therefore,

$$\begin{aligned} UBU^{CT} &= (C_{11}\dotplus\ldots\dotplus C_{1s}\dotplus C'_{11}\dotplus\ldots\dotplus C'_{1t}) + j(C_{21}\dotplus\ldots\dotplus C_{2s}\dotplus 0\dotplus\ldots\dotplus 0) \\ &= (C_{11}+jC_{21}) \dotplus \ldots \dotplus (C_{1s}+C_{2s}) \dotplus C'_{11} \dotplus \ldots \dotplus C'_{1t} \end{aligned}$$

where the C'_{1j} have only complex elements. Since UBU^{CT} is normal, so is each matrix in the above direct sum; there exist, then, unitary quaternion matrices W_k which diagonalize $C_{1k} + jC_{2k}$ and unitary complex matrices V_k which diagonalize C'_{1k}, for all the above k. If V is the unitary matrix formed by taking the appropriate direct sum of these W_k and V_k, it follows that $VUBU^{CT}V^{CT}$ is diagonal and that $VUAU^{CT}V^{CT} = VDV^{CT} = D$ is also diagonal. The converse is immediate.

The above generalizes as in the complex case:

THEOREM 6. *If $\{A_i\}$ is a set of normal quaternion matrices which commute in pairs, they can be diagonalized by the same unitary transformation.*

If each of the A_i have a single characteristic root, α_i, the theorem is true. If these roots are all real, the theorem is trivially true. If at least one root, say α_k, is non-real complex, let $VA_kV^{CT} = \alpha_k I$ and $VA_iV^{CT} = A'_i$ for all other i; then each A'_i commutes with $\alpha_k I$ and so all A'_i are normal, complex, and commutative in pairs, and can all be diagonalized by a complex unitary matrix U. Therefore, the unitary matrix UV diagonalizes all A_i.

In general, the proof follows by induction on the order of the A_i. The theorem is trivially true for 1×1 matrices. Assume the theorem to be true for $(n-1) \times (n-1)$ matrices. It may also be assumed that there is at least one matrix, A_j, which has at least two distinct roots; let $UA_jU^{CT} = D$ be diagonal (in the same form as D in the preceding theorem). Then each UA_iU^{CT} commutes with D, the problem is reduced to that involving matrices of order less than n and the theorem is true.

The following theorems are true in the complex case **(6)**; they are also true (obviously so from the isomorphism above) in the quaternion case:

THEOREM 7. *A quaternion matrix A is normal if and only if its polar matrices commute.*

THEOREM 8. *If A, B and AB are normal quaternion matrices, then BA is normal.*

THEOREM 9. *If A and B are normal quaternion matrices, then AB is normal if and only if each of A and B commutes with the hermitian polar matrix of the other.*

6. A diagonal form under unitary equivalence transformations. It is also possible to bring a quaternion matrix into a real diagonal matrix under a unitary equivalence transformation according to the following:

THEOREM 10. *For every $r \times s$ quaternion matrix A there exist two unitary quaternion matrices U and V (of dimensions $r \times r$ and $s \times s$, respectively) such that $UAV = D$ is diagonal with non-negative real roots along the diagonal.*

Let $A = A_1 + jA_2$ where A_1 and A_2 are complex, as before, but $r \times s$ in dimension. Let C be the $2r \times 2s$ matrix (composed of A_1 and A_2) with complex elements which corresponds to A. According to a corollary due to Eckert and Young **(2)**, if U is a $2r \times 2r$ unitary matrix which diagonalizes CC^{CT}, there exists a $2s \times 2s$ unitary matrix V such that $UCV = D_1$ is a $2r \times 2s$ diagonal matrix with non-negative real elements. From preceding work, this U may be taken as being in the form

$$\begin{bmatrix} U_1 & -U_2{}^C \\ U_2 & U_1{}^C \end{bmatrix},$$

so that $UCC^{CT}U^{CT} = D_2 \dotplus D_2$ is $2r \times 2r$ and so $UCV = D \dotplus D$ where D is $r \times s$, where the elements are non-negative real, and where $(D \dotplus D)(D \dotplus D)^{CT} = D_2 \dotplus D_2$. It remains to verify that V has the proper structure (i.e., like that of U). By considering the relation $UC = (D \dotplus D)V^{CT}$, it follows (as in the proof of the polar representation above) that V has this form where some arbitrariness may be involved, as before, in choosing V. If the components of V are V_1 and V_2, then $(U_1 + jU_2)A(V_1 + jV_2) = D$ as required in the theorem.

As in the complex case **(8)**, it is also true that

THEOREM 11. *If A and B are two $r \times s$ quaternion matrices, then there exist two unitary quaternion matrices U and V such that $UAV = D_1$ and $UBV = D_2$ are complex diagonal matrices if and only if AB^{CT} and $B^{CT}A$ are normal matrices.*

If such a U and V exist, the theorem is obviously true.

If AB^{CT} and $B^{CT}A$ are normal, but the preceding theorem $U_1AV_1 = D_1$ is a non-negative real diagonal matrix and $U_1BV_1 = C$. Let

$$D_1 = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad C_1 = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix},$$

where D is non-singular and B_1 has the same order as D. From the given conditions D_1C^{CT} and $C^{CT}D_1$ are normal; using the former, it follows that $(B_3D)(B_3D)^{CT} = 0$ (where B_3 has quaternion elements and D is real) so that $B_3D = 0$ and so $B_3 = 0$. Similarly $B_2 = 0$. Therefore $DB_1{}^{CT}$ and $B_1{}^{CT}D$ are normal. Now the characteristic roots of $DB_1{}^{CT}$ and $B_1{}^{CT}D$ are the same. (In the complex case, the characteristic roots of MN are the same as those of NM; from the isomorphism used above between $n \times n$ quaternion matrices and $2n \times 2n$ complex matrices, this result is seen to carry over). Therefore, from §5, there exists a polynomial $f(x)$ with real coefficients such that $B_1D = f(DB_1{}^{CT})$ and $DB_1 = f(B_1{}^{CT}D)$ and so $DB_1 = f(B_1{}^{CT}D) = D^{-1}f(DB_1{}^{CT})D = D^{-1}B_1DD$ or $D^2B_1 = B_1D^2$. Since D has positive diagonal elements, $DB_1 = B_1D$. Since $DB_1{}^{CT}B_1D = B_1D \cdot DB_1{}^{CT}$, then $B_1{}^{CT}B_1 = B_1B_1{}^{CT}$ and B_1 is a normal quaternion matrix which commutes with the (normal) real diagonal matrix D. There exists a quaternion unitary matrix W_1 which diagonalizes each simultaneously; there also exist unitary matrices W_2 and W_3 so that $W_2B_4W_3$ is a real diagonal matrix. By multiplying D_1 and C_1 each on the left and right, respectively, by the matrices

$$\begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix}, \quad \begin{bmatrix} W_1{}^{CT} & 0 \\ 0 & W_3 \end{bmatrix},$$

the theorem follows.

References

1. J. L. Brenner, *Matrices of quaternions*, Pac. J. Math., *1* (1951), 229–335.
2. C. Eckart and G. Young, *A principal axis transformation for non-hermitian matrices*, Bull. Amer. Math. Soc., *45* (1939), 118–121.
3. G. Frobenius, *Ueber lineare Substitutionen und bilineare Formen*, J. reine angew. Math., *84* (1878), 1–63.
4. H. C. Lee, *Eigenvalues and canonical forms of matrices with quaternion coefficients*, Proc. Royal Irish Academy, *52*, Section A, no. 117 (1949), 253–260.
5. H. Taber, *On the matrix equation* $\phi\Omega = \Omega\phi$, Proc. Amer. Acad. Arts. Sci., *26* (1890–1891), 64–66; *On a theorem of Sylvester's relating to non-degenerate matrices*, *27* (1891–1892), 46–56.
6. N. Wiegmann, *Normal products of matrices*, Duke Math. J., *15* (1948), 633–638.
7. J. Williamson, *A polar representation of singular matrices*, Bull. Amer. Math. Soc., *41* (1935), 118–123.
8. J. Williamson, *Note on a principal axis transformation for non-hermitian matrices*, Bull. Amer. Math., Soc., *45* (1939), 920–922.
9. A. Wintner and F. Murnaghan, *On a polar representation of non-singular square matrices*, Proc. Nat. Acad. Sci., U.S.A., *17* (1931), 676–678.
10. L. A. Wolf, *Similarity of matrices in which the elements are real quaternions*, Bull. Amer. Math. Soc., *42* (1936), 737–743.

Catholic University,
Washington, D.C.