A NORMAL FORM FOR A MATRIX UNDER THE UNITARY CONGRUENCE GROUP

D. C. YOULA

1. Introduction. Let C be a square matrix with complex elements. If $C = C'$ (C' denotes the transpose of C) there exists a unitary matrix U such that

$$(1) \qquad U'CU = \text{diag}[\mu_1, \mu_2, \ldots, \mu_n]$$

where the μ's are the non-negative square roots of the eigenvalues $\mu_1^2, \mu_2^2, \ldots, \mu_n^2$ of C^*C (C^* is the adjoint of C) **(2)**. If C is skew-symmetric, that is, $C = -C'$, there exists a unitary matrix U such that

$$(2) \qquad U'CU = \Sigma_1 \dotplus \Sigma_2 \dotplus \ldots \dotplus \Sigma_k \dotplus 0 \dotplus \ldots \dotplus 0$$

where

$$(3) \qquad \Sigma_r = \begin{bmatrix} 0 & \alpha_r \\ -\alpha_r & 0 \end{bmatrix}, \quad \alpha_r > 0, \quad r = 1, 2, \ldots, k,$$

and the α's are the positive square roots of the non-zero eigenvalues $\alpha_1^2, \alpha_2^2, \ldots, \alpha_k^2$ of C^*C **(1)**. Clearly rank $C = 2k$ and the number of zeros appearing in (2) is $n - 2k$. Both (1) and (2) are classical. In a recent paper **(3)** Stander and Wiegman, apparently unaware of **(1)**, give an alternative derivation of (2) with its appropriate generalization to quaternions.

The forms appearing on the right-hand sides of (1) and (2) possess the distinction of actually being canonic for symmetric and skew-symmetric matrices, respectively. The problem of finding a canonic form for an arbitrary matrix under the group of unitary congruence transformations is not only of mathematical interest but of the utmost importance for applied electrical engineering network theory, representing as it does the physical operation of embedding an n-port in a lossless, reciprocal "all-pass" $2n$-port **(4; 5)**. In this paper a start is made on this problem by obtaining a normal form for an arbitrary matrix under a U', U transformation and it is shown that (1) and (2) are immediate corollaries. Now $U'CU = W$, U unitary, implies that $U^*\bar{C}CU = \bar{W}W = U^{-1}\bar{C}CU$. Thus $\bar{C}C$ undergoes a similarity transformation and it is not surprising that a leading role is played by the eigenvalues of $\bar{C}C$. It is also shown that the normal form of C collapses into a direct sum of matrices if $\bar{C}C$ is normal. A set of sufficient conditions is given for this form to be canonic.

Received October 3, 1960. The work described in this memorandum was conducted at the Microwave Research Institute of the Polytechnic Institute of Brooklyn under Contract no. AF-19(604)-4143 with the Air Force Cambridge Research Center.

The technique employed is quite elementary and is inspired by an idea used by Schur in his treatment of the symmetric case **(2)**.

2. Notation and preliminary results. Let A be an arbitrary matrix. Then A', $\bar{A}$, A^*, A^{-1}, $d(A)$, and $r(A)$ denote the transpose, the complex conjugate, the adjoint $\bar{A}'$, the inverse, the determinant, and the rank of A respectively. Column vectors are indicated by $\mathbf{a}$, $\mathbf{b}$ etc., or in the alternative fashion $\mathbf{a} = (a_1, a_2, \ldots, a_n)'$ whenever it is desirable to exhibit the components explicitly. $\mathbf{0}_n$, $0_{m,n}$, and 1_n represent, in the same order, the n-dimensional zero vector, the $m \times n$ zero matrix, and $n \times n$ identity matrix.

Consider a general $n \times n$ matrix of complex elements $C = B + iD$ where B and D are its real and imaginary parts. To C can be assigned the real $2n \times 2n$ matrix

$$T(C) = \left[\begin{array}{c|c} B & D \\ \hline D & -B \end{array}\right]. \tag{4}$$

The properties of $T(C)$ are subsumed in Lemma 1 **(2)**.

LEMMA 1. (1) *The eigenvalues of $T(C)$ occur in negative pairs.* (2) *The eigenvalues of $T(C)$ occur in complex conjugate pairs.* (3) *Let the matrix $\bar{C}C$ possess the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_\nu$ with respective multiplicities $r_1, r_2, \ldots, r_\nu$. Then the distinct eigenvalues of $T^2(C)$ are $\lambda_1, \lambda_2, \ldots, \lambda_\nu$ with respective multiplicities $2r_1, 2r_2, \ldots, 2r_\nu$.*

Proof. By direct verification

$$\sigma_1 T(C)\sigma_1^{-1} = -T(C) \tag{5}$$

where

$$\sigma_1 = \left[\begin{array}{c|c} 0 & 1_n \\ \hline -1_n & 0 \end{array}\right], \qquad \sigma_1^2 = -1_{2n}. \tag{6}$$

(2) Since $T(C)$ is real, its characteristic equation has real coefficients. (3) A straightforward calculation yields

$$\sigma_2 T^2(C)\sigma_2^{-1} = \bar{C}C \dotplus C\bar{C}$$

where

$$\sigma_2 = \frac{1}{\sqrt{2}}\left[\begin{array}{c|c} 1_n & -i1_n \\ \hline -i1_n & 1_n \end{array}\right], \qquad \sigma_2\sigma_2^* = 1_{2n},$$

Q.E.D.

LEMMA 2. *Let $C = B + iD$ (B and D real) be an $n \times n$ matrix. There exists a non-trivial n-vector $\mathbf{a}$ and a scalar ξ such that*

$$C\mathbf{a} = \xi\bar{\mathbf{a}}$$

if and only if $\bar{C}C$ possesses a non-negative eigenvalue.

Proof. Suppose (9) is satisfied for some non-trivial $\mathbf{a}$. Then,

$$\bar{C}C\mathbf{a} = \xi\bar{C}\bar{\mathbf{a}} = |\xi|^2\mathbf{a},$$

that is, $\mathbf{a}$ is an eigenvector of $\bar{C}C$ corresponding to the non-negative eigenvalue $|\xi|^2$. Conversely if $\bar{C}C$ has an eigenvalue $\xi^2 \geqslant 0$, $\xi \geqslant 0$, it follows, by Lemma 1, parts (1) and (3), that $\pm\,\xi$ are both eigenvalues of $T(C)$. Thus there exists a real $2n$-vector

$$\psi = \begin{array}{c} \left[\begin{array}{c} \mathbf{a}_1 \\ \hline -\mathbf{a}_2 \end{array}\right] \end{array}\begin{array}{l} n \\ n \end{array} \tag{7}$$

such that

$$\left[\begin{array}{c|c} B & D \\ \hline D & -B \end{array}\right]\psi = \xi\psi, \tag{8}$$

or, in expanded form

$$\xi\mathbf{a}_1 = B\mathbf{a}_1 - D\mathbf{a}_2,$$
$$-\xi\mathbf{a}_2 = D\mathbf{a}_1 + B\mathbf{a}_2.$$
$$\therefore \xi(\mathbf{a}_1 - i\mathbf{a}_2) = (B + iD)(\mathbf{a}_1 + i\mathbf{a}_2).$$
$$\therefore \qquad \xi\bar{\mathbf{a}} = C\mathbf{a}, \tag{9}$$
$$\mathbf{0} \neq \mathbf{a} = \mathbf{a}_1 + i\mathbf{a}_2, \tag{10}$$

Q.E.D.

LEMMA 3. *Let $\bar{C}C$ have the non-negative eigenvalues $\xi_1^2, \xi_2^2, \ldots, \xi_\nu^2$, $\xi_r \geqslant 0$, $r = 1, 2, \ldots, \nu$. Then there exists a unitary matrix U such that*

$$U'CU = \begin{array}{c} \begin{array}{cc} \nu & n-\nu \end{array} \\ \left[\begin{array}{c|c} \Delta & C_1 \\ \hline O & C_2 \end{array}\right] \end{array}\begin{array}{l} \\ \nu \\ n-\nu \end{array} \tag{11}$$

where Δ is a $\nu \times \nu$ upper triangular matrix with diagonal elements $\xi_1, \xi_2, \ldots, \xi_\nu$.

Proof. According to Lemma 2 there exists a non-zero n-vector $\mathbf{a}_1$ such that $\mathbf{a}_1^*\mathbf{a}_1 = 1$ and $C\mathbf{a}_1 = \xi_1\bar{\mathbf{a}}_1$. Let A be an $n \times n$ unitary matrix with columns $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. Then,

$$\begin{aligned} (A'CA)_{k,1} &= \mathbf{a}_k'C\mathbf{a}_1 = \xi_1\mathbf{a}_k'\bar{\mathbf{a}}_1 \\ &= 1, \qquad k = 1, \\ &= 0, \qquad k > 1. \end{aligned}$$

Hence

$$A'CA = \begin{array}{c} \left[\begin{array}{c|c} \xi_1 & M \\ \hline O & N \end{array}\right] \\ \begin{array}{cc} 1 & n-1 \end{array} \end{array}\begin{array}{l} 1 \\ n-1 \\ \\ \end{array} \tag{12}$$

and

$$A^*\bar{C}CA = \begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} \xi_1^2 & P \\ \hline O & \bar{N}N \end{array}\right] \end{array}\begin{array}{l} \\ 1 \\ n-1' \end{array} \tag{13}$$

https://doi.org/10.4153/CJM-1961-059-8 Published online by Cambridge University Press

where

$$P = \xi_1 M + \bar{M}N. \tag{14}$$

Now N is an $(n - 1) \times (n - 1)$ matrix and since the remaining eigenvalues of $\bar{C}C$ and $\bar{N}N$ coincide, the proof of the lemma is very easily completed by induction. Let V be an $(n - 1) \times (n - 1)$ unitary matrix such that

$$V'NV = \begin{array}{cc} \begin{array}{cc} \nu - 1 & n - \nu \end{array} & \\ \left[\begin{array}{c|c} \delta & \hat{M} \\ \hline 0 & C_2 \end{array}\right] & \begin{array}{l} \nu - 1 \\ n - \nu \end{array} \end{array}$$

where δ is a $(\nu - 1) \times (\nu - 1)$ upper triangular matrix with the diagonal entries $\xi_2, \xi_3, \ldots, \xi_\nu$. Then, the $n \times n$ matrix

$$U \equiv A\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & V \end{array}\right]$$

is unitary and

$$U'CU = \begin{array}{cc} \left[\begin{array}{c|c|c} \xi_1 & E & F \\ \hline 0 & \delta & \hat{M} \\ \hline 0 & 0 & C_2 \end{array}\right] & \begin{array}{l} 1 \\ \nu - 1 \\ n - \nu \end{array} \end{array}$$

$$= \left[\begin{array}{c|c} \Delta & C_1 \\ \hline 0 & C_2 \end{array}\right]$$

in which

$$MV = [E|F],$$

$$\Delta = \begin{array}{cc} \left[\begin{array}{c|c} \xi_1 & E \\ \hline 0 & \delta \end{array}\right] & \begin{array}{l} 1 \\ \nu - 1 \end{array} \end{array}$$

and

$$C_1 = \left[\begin{array}{c} F \\ \hline \hat{M} \end{array}\right],$$

Q.E.D.

By applying Lemma 3 to C' it is easily established that there is also a unitary matrix V such that

$$V'CV = \begin{array}{cc} \left[\begin{array}{c|c} \hat{\Delta} & O \\ \hline Q & R \end{array}\right] & \begin{array}{l} \nu \\ n - \nu \end{array} \\ \begin{array}{cc} \nu & n - \nu \end{array} & \end{array}, \tag{15}$$

$\hat{\Delta}$ being a lower triangular $\nu \times \nu$ matrix whose diagonal elements are $\xi_1, \xi_2, \ldots, \xi_\nu$.

LEMMA 4. *Let* $-\lambda^2 \ngeq 0$ *be an eigenvalue of* $\bar{C}C$. *There exist two linearly independent n-vectors* **a** *and* **b** *such that*

(16) $$C\mathbf{a} = \lambda\bar{\mathbf{b}}$$
and

(17) $$C\mathbf{b} = -\lambda\bar{\mathbf{a}}.$$

Proof. By hypothesis $-\lambda^2 \ngtr 0$ is an eigenvalue of $\bar{C}C$. Let $\mathbf{a}$ be an associated eigenvector:

(18) $$\bar{C}C\mathbf{a} = -\lambda^2\mathbf{a}.$$
Define the vector $\mathbf{b}$ by

(19) $$C\mathbf{a} = \lambda\bar{\mathbf{b}}.$$
From (19) and (18)

$$\bar{C}C\mathbf{a} = \lambda\bar{C}\bar{\mathbf{b}} = -\lambda^2\mathbf{a}\,.$$
$$\therefore C\mathbf{b} = -\bar{\lambda}\bar{\mathbf{a}}.$$

Now suppose that $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent, that is, $\mathbf{b} = \mu\mathbf{a}$, μ a scalar. Substituting in (19), $C\mathbf{a} = \lambda\bar{\mu}\bar{\mathbf{a}}$, whence, using (19) once again,

(20) $$\bar{C}C\mathbf{a} = \lambda\bar{\mu}\bar{C}\bar{\mathbf{a}} = |\lambda\mu|^2\mathbf{a}.$$

Equation (20) shows that $\mathbf{a}$ is also an eigenvector of $\bar{C}C$ corresponding to the positive eigenvalue $|\lambda\mu|^2 \neq -\lambda^2$. But this is impossible because eigenvectors belonging to distinct characteristic values must be linearly independent, Q.E.D.

Since $\bar{C}C\mathbf{b} = -(\bar{\lambda})^2\mathbf{b}$, $\mathbf{b}$ is also an eigenvector of $\bar{C}C$ with eigenvalue $-(\bar{\lambda})^2$. If λ is real $\mathbf{a}$ and $\mathbf{b}$ are two linearly independent eigenvectors of $\bar{C}C$ corresponding to the same negative eigenvalue $-\lambda^2 < 0$. Hence any negative eigenvalue of a matrix of the form $\bar{C}C$ is at least of algebraic multiplicity two. It will soon appear that its multiplicity is always an even integer. More important is the observation that for $-\lambda^2 < 0$, $\lambda > 0$, it may always be assumed, without loss of generality, that $\mathbf{a}^*\mathbf{b} = 0$. For, if $C\mathbf{a}_0 = \lambda\bar{\mathbf{b}}_0$ and $C\mathbf{b}_0 = -\lambda\bar{\mathbf{a}}_0$, $C\mathbf{x} = \lambda\bar{\mathbf{y}}$, where $\mathbf{x} = \mathbf{a}_0 + \mu\mathbf{b}_0$, $\mathbf{y} = \mathbf{b}_0 - \bar{\mu}\mathbf{a}_0$ and μ is an arbitrary scalar. By choosing μ to be one of the roots of the quadratic equation

$$(\overset{*}{\mathbf{a}}_0\mathbf{b}_0)\mu^2 + (\overset{*}{\mathbf{a}}_0\mathbf{a}_0 - \overset{*}{\mathbf{b}}_0\mathbf{b}_0)\mu - \overline{\overset{*}{\mathbf{a}}_0\mathbf{b}_0} = 0,$$

$\mathbf{x}$ can be made orthogonal to $\mathbf{y}$. Since $\mathbf{x}$ is also an eigenvector of $\bar{C}C$ associated with the eigenvalue $-\lambda^2 < 0$, the result follows.

If λ is complex and $-\lambda^2 \ngtr 0$, one resorts to the Schmidt process to create an orthonormal pair $\mathbf{x}$ and $\mathbf{y}$ (choose $\mathbf{a}^*\mathbf{a} = 1$ from the outset). Let

(21) $$\mathbf{a} = \mathbf{x}$$
and

(22) $$\mathbf{b} = \mu\mathbf{x} + \eta\mathbf{y},$$

μ and η being scalars. The conditions $\mathbf{x}^*\mathbf{y} = 0$, $\mathbf{x}^*\mathbf{x} = 1$ yield

(23) $$\mu = \mathbf{a}^*\mathbf{b},$$

(24) $$|\eta|^2 = (\mathbf{b} - \mu\mathbf{a})^*(\mathbf{b} - \mu\mathbf{a}) > 0.$$

https://doi.org/10.4153/CJM-1961-059-8 Published online by Cambridge University Press

Equation (24) defines only the magnitude of η which may therefore be chosen real and positive. Similarly, by changing the argument of $\mathbf{a}$ suitably, $\mu \geqslant 0$. As has already been pointed out, if $\lambda > 0$, $\mu = 0$. Substitution of (21) and (22) into (16) and (17) gives

$$\mathbf{x}'C\mathbf{x} = \lambda\mu, \tag{25}$$

$$\mathbf{y}'C\mathbf{x} = \lambda\eta, \tag{26}$$

$$\mathbf{x}'C\mathbf{y} = -\frac{1}{\eta}(\bar{\lambda} + \lambda\mu^2), \tag{27}$$

$$\mathbf{y}'C\mathbf{y} = -\lambda\mu, \tag{28}$$

$\eta > 0$, $\mu \geqslant 0$.

Let A be an $n \times n$ unitary matrix with $\mathbf{x}$ and $\mathbf{y}$ for its first two columns. Then, by equations (25)–(28),

$$A'CA = \begin{array}{cc} \begin{array}{cc} 2 & n-2 \end{array} & \\ \left[\begin{array}{c|c} \Sigma & C_1 \\ \hline O & C_2 \end{array}\right] & \begin{array}{l} 2 \\ n-2 \end{array} \end{array} \tag{29}$$

where

$$\Sigma = \left[\begin{array}{c|c} \lambda\mu & -\frac{1}{\eta}(\bar{\lambda} + \lambda\mu^2) \\ \hline \lambda\eta & -\lambda\mu \end{array}\right], \qquad \mu \geqslant 0,\ \eta > 0. \tag{30}$$

If $\lambda > 0$, $\mu = 0$ and

$$\Sigma = \left[\begin{array}{c|c} O & -\frac{\lambda}{\eta} \\ \hline \lambda\eta & O \end{array}\right], \qquad \lambda > 0,\ \eta > 0. \tag{31}$$

Observe that

$$A^*\bar{C}CA = \left[\begin{array}{c|c} \bar{\Sigma}\,\Sigma & \bar{\Sigma}\,C_1 + \bar{C}_1C_2 \\ \hline O & \bar{C}_2C_2 \end{array}\right] \tag{31a}$$

and

$$\bar{\Sigma}\,\Sigma = \left[\begin{array}{c|c} -\lambda^2 & \frac{\mu}{\eta}(\lambda^2 - \bar{\lambda}^2) \\ \hline O & -(\bar{\lambda})^2 \end{array}\right]. \tag{31b}$$

A repetition of the process on C_2, etc., leads to the conclusion that any negative eigenvalue of $\bar{C}C$ is of even multiplicity because each successive step contributes two identical ones. Combining this with the fact that the eigenvalues of $\bar{C}C$ occur in complex conjugate pairs leads to Lemma 5.

LEMMA 5. *Let C be an arbitrary $n \times n$ matrix. Those eigenvalues $-\lambda^2$ of $\bar{C}C$ which do not lie in the half-line $-\lambda^2 \geqslant 0$ may be arranged in the tableau*

$$\begin{array}{l} -\lambda_1^2,\ -\lambda_2^2, \ldots, -\lambda_k^2, \\ -(\bar{\lambda}_1)^2,\ -(\bar{\lambda}_2)^2, \ldots, -(\bar{\lambda}_k)^2 \end{array} \tag{32}$$

and there exists a unitary matrix U such that

$$U'CU = \begin{array}{c} \left[\begin{array}{c|c} \Omega & C_1 \\ \hline O & C_2 \end{array}\right] \begin{array}{l} 2k \\ n-2k \end{array} \\ \begin{array}{cc} 2k & n-2k \end{array} \end{array} \tag{33}$$

where

$$\Omega = \begin{array}{c} \begin{bmatrix} \Sigma_1 & x & x & x \\ 0 & \Sigma_2 & x & x \\ 0 & 0 & \ddots & x \\ 0 & 0 & 0 & \Sigma_k \end{bmatrix} 2k \\ 2k \end{array} \tag{34}$$

is a block upper triangular matrix in which the 2×2 *Σ's are defined in* (30) *and the crosses indicate the possible presence of non-zero* 2×2 *blocks. Observe that* (31b) *implies that* $\bar{\Omega}\Omega$ *is upper triangular.*

Proof. Induction.

Enough material is now on hand for the main theorem.

3. A normal form for a matrix under the unitary congruence group.

THEOREM 1. *Let C be an arbitrary $n \times n$ matrix. The eigenvalues of $\bar{C}C$ may be arranged as follows:*

$$\begin{array}{ll} \xi_1^2, \xi_2^2, \ldots, \xi_\nu^2, \xi_r \geqslant 0, & r = 1, 2, \ldots, \nu, \\ -\lambda_1^2, -\lambda_2^2, \ldots, -\lambda_k^2, -\lambda_j^2 \not\geq 0, & j = 1, 2, \ldots, k, \\ -(\bar{\lambda}_1)^2, -(\bar{\lambda}_2)^2, \ldots, -(\bar{\lambda}_k)^2, & \nu + 2k = n. \end{array} \tag{35}$$

Furthermore, there exists an $n \times n$ unitary matrix U such that

$$U'CU = \begin{array}{c} \left[\begin{array}{c|c} \Delta & C_1 \\ \hline 0 & \Omega \end{array}\right] \begin{array}{l} \nu \\ 2k \end{array} \\ \begin{array}{cc} \nu & 2k \end{array} \end{array} \tag{36}$$

where Δ is a $\nu \times \nu$ upper triangular matrix with diagonal elements $\xi_1, \xi_2, \ldots, \xi_\nu$ and Ω is a $2k \times 2k$ block upper triangular matrix possessing the structure (34) *in which the* 2×2 *Σ's are as defined in* (30) *for complex λ and* (31) *for $\lambda > 0$.*

Proof. Lemmas 3 and 5.

COROLLARY 1. *Let $C = C'$ be a symmetric $n \times n$ matrix. There exists an $n \times n$ unitary matrix U which reduces C to the diagonal form* (1).

Proof. Since $\bar{C}C = C^*C$, all eigenvalues of $\bar{C}C$ are non-negative. Hence $\nu = n$ and $k = 0$. By Theorem 1 there exists a unitary matrix U such that $U'CU = \Delta$, an $n \times n$ upper triangular matrix. But Δ must be symmetric and therefore is actually diagonal, Q.E.D.

COROLLARY 2. *Let $C = -C'$ be an arbitrary $n \times n$ skew-symmetric matrix. There exists an $n \times n$ unitary matrix U which reduces C to the direct sum form* (2).

Proof. Since $\bar{C}C = -C^*C$, all eigenvalues of $\bar{C}C$ are either negative or zero. Let U be that unitary matrix which transforms C into the right-hand side of (36). The skew-symmetry of C implies that $\Delta = 0_{\nu,\nu}$, $C_1 = 0_{\nu,2k}$, and Ω is the direct sum

$$\Omega = \Sigma_1 \dotplus \Sigma_2 \dotplus \ldots \dotplus \Sigma_k,$$

each Σ having the structure (31). Now the Σ's themselves must also be skew-symmetric and therefore all η's $= 1$. That is to say,

$$\Sigma_r = \left[\begin{array}{c|c} 0 & -\lambda_r \\ \hline \lambda_r & 0 \end{array}\right], \lambda_r > 0, \qquad r = 1, 2, \ldots, k,$$

Q.E.D.

COROLLARY 3. *Let C be an $n \times n$ matrix. There exists an $n \times n$ unitary matrix U such that $U'CU$ is either upper or lower triangular if and only if the eigenvalues of $\bar{C}C$ are all non-negative.*

COROLLARY 4. *Let C be an $n \times n$ matrix. There exists an $n \times n$ unitary matrix U such that $U'CU$ is a block upper triangular matrix Ω (see* (34)), *the Σ's being defined by* (31), *if and only if all eigenvalues of $\bar{C}C$ are negative.*

In either of the two cases C symmetric or C skew-symmetric, $\bar{C}C$ is normal and it is reasonable to expect that (36) simplifies considerably even under this more general restriction. A complete answer is available and easily obtained with the aid of Lemma 6.

LEMMA 6. *The matrix*

$$Q = \begin{array}{cc} \left[\begin{array}{c|c} A & B \\ \hline O & C \end{array}\right] & \begin{array}{c} k \\ r \end{array} \\ \begin{array}{cc} k & r \end{array} & \end{array} \tag{37}$$

is normal if and only if $B = O_{k,r}$ and A and C are normal.

Proof. The condition $QQ^* = Q^*Q$ yields the three matrix equations

$$A^*A - AA^* = BB^*, \tag{38}$$

$$A^*B = BC^*, \tag{39}$$

$$CC^* - C^*C = B^*B. \tag{40}$$

By (38), trace$(BB^*) = 0$. Thus $B = O_{k,r}$, $A^*A = AA^*$, and $C^*C = CC^*$, Q.E.D.

One corollary of this lemma is that any normal block-upper or lower triangular matrix Ω (the Σ's can now be arbitrary square blocks) is a direct sum $\Omega = \Sigma_1 \dot{+} \Sigma_2 \dot{+} \ldots \dot{+} \Sigma_k$.

THEOREM 2. *Let C be an $n \times n$ matrix. There exists an $n \times n$ unitary matrix U such that*

$$U'CU = \Delta \dot{+} \Omega \tag{41}$$

where

(1) *Δ is square and upper triangular,*

(2) *$\bar{\Delta}\Delta$ is diagonal,*

(3) *Ω is a block upper triangular matrix of the form* (34), *the blocks Σ being defined generically by*

$$\Sigma = \left[\begin{array}{c|c} 0 & \frac{\bar{\lambda}}{\eta} \\ \hline \eta\lambda & 0 \end{array}\right], \quad \eta > 0, \tag{41a}$$

and

(4) *$\bar{\Omega}\Omega$ is diagonal,*

if and only if $\bar{C}C$ is normal.

Proof. Suppose such a U exists, Δ and Ω possessing the properties enumerated in (1) to (4). Then,

$$U^*\bar{C}CU = \bar{\Delta}\Delta \dot{+} \bar{\Omega}\Omega,$$

a diagonal matrix. Since $\bar{C}C$ is unitarily similar to a diagonal matrix it is normal.

Conversely, let $\bar{C}C$ be normal. By Theorem 1 there exists an $n \times n$ unitary matrix U such that

$$U'CU = \begin{array}{c} \left[\begin{array}{c|c} \Delta & C_1 \\ \hline O & \Omega \end{array}\right]\begin{array}{l} \nu \\ 2k \end{array} \\ \begin{array}{cc} \nu & 2k \end{array} \end{array} \tag{42}$$

where Δ is upper triangular and Ω is as described in (31) and (34). Now

$$\overset{*}{U}\bar{C}CU = \begin{array}{c} \left[\begin{array}{c|c} \bar{\Delta}\Delta & \Delta C_1 + \bar{C}_1\Omega \\ \hline O & \bar{\Omega}\Omega \end{array}\right]\begin{array}{l} \nu \\ 2k \end{array} \\ \begin{array}{cc} \nu & 2k \end{array} \end{array}. \tag{43}$$

The normality of $\bar{C}C$ implies that of the right-hand side of (43). By Lemma 6,

$$\begin{aligned} &\text{(a)}\ \ \Delta C_1 + \bar{C}_1\Omega = O_{\nu,2k} \\ &\text{(b)}\ \ \bar{\Delta}\Delta \text{ is normal;} \\ &\text{(c)}\ \ \bar{\Omega}\Omega \text{ is normal.} \end{aligned} \tag{44}$$

But $\bar{\Delta}\Delta$ and $\bar{\Omega}\Omega$ are already upper triangular, and must therefore be diagonal. According to (44),

https://doi.org/10.4153/CJM-1961-059-8 Published online by Cambridge University Press

$$\Delta C_1 = -\bar{C}_1\Omega$$

or

(45) $$\bar{\Delta}\Delta C_1 = -\bar{\Delta}\bar{C}_1\Omega = C_1\bar{\Omega}\Omega.$$

However, the eigenvalues of $\bar{\Delta}\Delta$ are all non-negative and those of $\bar{\Omega}\Omega$ are either negative or complex. Thus $\bar{\Delta}\Delta$ and $\bar{\Omega}\Omega$ have no eigenvalues in common and by a well-known theorem the only solution of (44) is $C_1 = O_{\nu,2k}$. As regards (41a), refer to equations (23) and (24). Since $\bar{C}C$ is normal any two eigenvectors **a** and **b** corresponding to the complex characteristic values $-\lambda^2$ and $(-\bar{\lambda})^2$ must be orthogonal and therefore all μ's are zero and (30) collapses into (41a), Q.E.D.

COROLLARY. *A sufficient set of conditions for* (41) *to be canonic* (*under a prescribed ordering*) *is the following: any negative eigenvalue of* $\bar{C}C$ *is of algebraic multiplicity not exceeding two and all other eigenvalues are simple.*

Proof. Let $U'CU = \Delta \dotplus \Omega$ where (see Theorem 2 and (31b))

(46) $$\bar{\Delta}\Delta = \operatorname{diag}[\xi_1^2, \xi_2^2, \ldots, \xi_\nu^2] = D_1, \quad \xi_i \geqslant 0, \qquad (i = 1, 2, \ldots, \nu),$$

and

(47) $$\bar{\Omega}\Omega = -\operatorname{diag}[\lambda_1^2, (\bar{\lambda}_1)^2, \lambda_2^2, (\bar{\lambda}_2)^2, \ldots, \lambda_k^2, \quad (\bar{\lambda}_k)^2] = -D_2.$$

From (46), $\Delta\bar{\Delta}\Delta = \Delta D_1 = \bar{D}_1\Delta = D_1\Delta$. Thus,

(48) $$(\Delta)_{rj}\xi_j^2 = (\Delta)_{rj}\xi_r^2, \qquad (r, j = 1, 2, \ldots, \nu).$$

Since, by hypothesis, $\xi_r^2 \neq \xi_j^2$, $r \neq j$, $(\Delta)_{rj} = 0$, $r \neq j$ and $\Delta = \operatorname{diag}[\xi_1, \xi_2, \ldots, \xi_\nu]$.

Again, from (47), $-\Omega\bar{\Omega}\Omega = \Omega D_2 = \bar{D}_2\Omega$. Now Ω is of even order $2k$ and may be partitioned into 2×2 blocks. For example, if $2k = 6$

(49) $$\Omega = \begin{array}{cc} \begin{array}{ccc} 2 & 2 & 2 \end{array} & \\ \begin{bmatrix} \Sigma_1 & P & Q \\ O & \Sigma_2 & R \\ O & O & \Sigma_3 \end{bmatrix} & \begin{array}{c} 2 \\ 2 \\ 2 \end{array} \end{array}.$$

The matrix equation $\Omega D_2 = \bar{D}_2\Omega$ gives

$$P \operatorname{diag}[\lambda_2^2, (\bar{\lambda}_2)^2] = \operatorname{diag}[(\bar{\lambda}_1)^2, \lambda_1^2]P.$$

From the assumptions, $\bar{\lambda}_2^2 \neq \lambda_1^2 \neq \lambda_2^2$ and therefore $P = O_{2,2}$. All other 2×2 blocks are treated in the same way and shown to be zero matrices. Consequently

(50) $$U'CU = \xi_1 \dotplus \xi_2 \dotplus \ldots \dotplus \xi_\nu \dotplus \Sigma_1 \dotplus \Sigma_2 \dotplus \ldots \dotplus \Sigma_k$$

where

(51) $$\Sigma_r = \left[\begin{array}{c|c} O & \dfrac{\bar{\lambda}_r}{\eta_r} \\ \hline -\eta_r\lambda_r & O \end{array}\right], \qquad (r = 1, 2, \ldots, k).$$

All that remains is the evaluation of the η's. From (50) and (51),

$$U^{-1}C^*CU = \text{diag}[\xi_1^2, \xi_2^2, \ldots, \xi_\nu^2, \eta_1^2|\lambda_1|^2, \eta_1^{-2}|\lambda_1|^2, \ldots, \eta_k^2|\lambda_k|^2, \eta_k^{-2}|\lambda_k|^2]. \tag{52}$$

Let the ξ's be ordered so that $\xi_1 \leqslant \xi_2 \leqslant, \ldots, \leqslant \xi_\nu$ and the λ's so that $|\lambda_1| \leqslant |\lambda_2| \leqslant, \ldots, \leqslant |\lambda_k|$. Assume the eigenvalues of C^*C to be ordered according to the scheme

$$\xi_1^2, \xi_2^2, \ldots, \xi_\nu^2, \alpha_{11}^2, \alpha_{12}^2, \alpha_{21}^2, \alpha_{22}^2, \ldots, \alpha_{k1}^2, \alpha_{k2}^2$$

in which all α's > 0 and

$$\alpha_{11}\alpha_{12} \leqslant \alpha_{21}\alpha_{22} \leqslant, \ldots, \leqslant \alpha_{k1}\alpha_{k2}.$$

Then,

$$\eta_r = \sqrt{\left(\frac{\alpha_{r1}}{\alpha_{r2}}\right)}, \qquad (r = 1, 2, \ldots, k). \tag{53}$$

In other words the η's are uniquely determined by C, Q.E.D.

The reader shuold note that the above corollary does not necessarily include the cases $C = C'$ or $C = -C'$ which in general require separate arguments.

References

1. L. K. Hua, *On the theory of automorphic functions of a matrix variable I-geometrical basis*, Amer. J. Math., *66* (1944).
2. I. Schur, *Ein Satz ueber quadratische Formen mit komplexen Koeffizienten*, Amer. J. Math., *67* (1942), 472.
3. J. W. Stander and N. A. Wiegman, *Canonical forms for certain matrices under unitary congruence*, Can. J. Math., *12* (1960).
4. D. C. Youla, *Direct single frequency synthesis from a prescribed scattering matrix*, IRE Transactions of the Professional Group on Circuit Theory, vol. CT-6, no. 4 (December, 1959).
5. ——— *Some new and useful matrix results*, Mem. 34, Polytechnic Institute of Brooklyn, Microwave Research Institute (March 7, 1960).

Polytechnic Institute of Brooklyn