Can. J. Math., Vol. XXIV, No. 4, 1972, pp. 686–695

RATIONAL TENSOR REPRESENTATIONS OF Hom(V, V) AND AN EXTENSION OF AN INEQUALITY OF I. SCHUR.

MARVIN MARCUS AND WILLIAM ROBERT GORDON

1. Introduction. Let V be an n-dimensional vector space over the complex numbers equipped with an inner product (x, y), and let (P, μ) be a symmetry class in the mth tensor product of V associated with a permutation group G and a character χ (see below). Then for each $T \in \mathrm{Hom}(V, V)$ the function φ which sends each m-tuple $(v_1, \ldots, v_m)$ of elements of V to the tensor $\mu(Tv_1, \ldots, Tv_m)$ is symmetric with respect to G and χ, and so there is a unique linear map $K(T)$ from P to P such that $\varphi = K(T)\mu$.

It is easily checked that K: $\mathrm{Hom}(V, V) \to \mathrm{Hom}(P, P)$ is a rational representation of the multiplicative semi-group in $\mathrm{Hom}(V, V)$: for any two linear operators S and T on V

$$K(ST) = K(S)K(T).$$

Moreover, if T is normal then, with respect to the inner product induced on P by the inner product on V (see below), $K(T)$ is normal.

In this paper we prove

THEOREM 1. *If S and T are in* $\mathrm{Hom}(V, V)$ *and* rank $T > m$, *then* $K(T) = K(S)$ *if and only if $T = cS$ for some m*th *root of unity, c.*

THEOREM 2. *If $T \in Hom(V, V)$ and* rank $T > m$, *then $K(T)$ is normal if and only if T is normal.*

By considering an $n \times n$ complex matrix as a linear operator on complex n-tuple space, we have

THEOREM 3. *If A is an $n \times n$ complex matrix with* rank $A = m$ *and if $K(A)$ is normal, then A is unitarily similar to the direct sum of a non-singular $m \times m$ upper triangular matrix and the $(n - m) \times (n - m)$ zero matrix.*

We shall show in § 4 how these results can be easily applied to produce the following interesting theorem which was announced recently by R. Kess, H. L. de Vries, and R. Wegmann [**1**].

THEOREM 4. *If A is a non-normal $n \times n$ complex matrix with eigenvalues*

Received August 4, 1971 and in revised form, November 16, 1971. The research of both authors was supported by the U. S. Air Force Office of Scientific Research under grant AFOSR 698–67.

$\lambda_1, \ldots, \lambda_n$, *if* $D = AA^* - A^*A$, *and if* $\|\ \|$ *denotes the usual Euclidean matrix norm, then*

$$\sum_{i=1}^{n} |\lambda_i|^2 \leqq (\|A\|^4 - \tfrac{1}{2}\|D\|^2)^{\frac{1}{2}} \tag{1}$$

with equality if and only if

$$A = \alpha(vw^* + rwv^*), \tag{2}$$

where α *is a non-zero complex number,* r *is a real number,* $0 \leqq r < 1$, *and where* v *and* w *are orthonormal complex n-tuples.*

It will be seen from our proof of this theorem that inequality (1) is an application of Schur's well known inequality [**2**] to the appropriate transformations.

2. Definitions and notation. Throughout this paper, V will be a finite-dimensional inner product space over the complex numbers **C**, dim $V = n$, G a subgroup of S_m, the symmetric group of degree m, and χ a character of degree 1 on G, i.e., a homomorphism of G into the unit circle. If V is a vector space over **C**, and $\varphi(v_1, \ldots, v_m)$ is an m-multilinear function on the cartesian product $\times_1^m V$ to U, then φ is said to be symmetric with respect to G and χ if

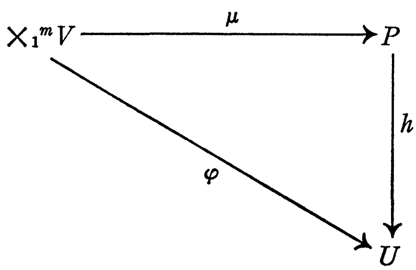
$$\varphi(v_{\sigma(1)}, \ldots, v_{\sigma(m)}) = \chi(\sigma)\varphi(v_1, \ldots, v_m),$$

for any $\sigma \in G$ and for arbitrary $v_i \in V$. By a symmetry class of tensors over V associated with G and χ we shall mean a pair (P, μ), consisting of a vector space P over **C** and an m-multilinear function μ: $\times_1^m V \rightarrow P$, symmetric with respect to G and χ, which is universal for these properties; that is;

(i) $\langle \text{rng } \mu \rangle = P$; i.e., the linear closure of the range of μ is P.

(ii) (Universal Factorization Property) For any vector space U over **C** and any m-multilinear function φ: $\times_1^m V \rightarrow U$, symmetric with respect to G and χ, there exists a linear h: $P \rightarrow U$ such that $\varphi = h\mu$.

(3)



The symmetry class (P, μ) is unique to within canonical isomorphisms, and the linear map h is uniquely determined by φ. The element $\mu(v_1, \ldots, v_m) \in P$ is called decomposable and will sometimes be denoted by $v_1 * \ldots * v_m$. The three most familiar symmetry classes are: (i) the space of m-contravariant tensors, $P = \otimes_1^m V$, $\mu(v_1, \ldots, v_m) = v_1 \otimes \ldots \otimes v_m$, i.e., $G = \{e\}$; (ii) the mth exterior power of V, $P = \wedge^m V$, $\mu(v_1, \ldots, v_m) = v_1 \wedge \ldots \wedge v_m$, i.e., $G = S_m$ and $\chi = \text{sgn} = \epsilon$; (iii) the mth completely symmetric space over V, $P = V^{(m)}$, $\mu(v_1, \ldots, v_m) = v_1 \ldots v_m$, i.e., $G = S_m$ and $\chi \equiv 1$.

Any symmetry class of tensors (P, μ) can be realized as a subspace of $\otimes_1^m V$ by defining

$$\mu(v_1, \ldots, v_m) = \frac{1}{|G|} \sum_{\sigma \in G} \chi(\sigma) v_{\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma^{-1}(m)}.$$

In order to describe a basis for an arbitrary symmetry class associated with G and χ, we regard the elements of G as permutations acting on the set of all sequences of length m chosen from the integers $1, \ldots, n$. That is, $\Gamma_n^m = Z_n^{Z_m}$, where $Z_m = \{1, \ldots, m\}$ and for $\sigma \in G$, $\gamma \in \Gamma_n^m$

$$\sigma(\gamma)(t) = \gamma(\sigma^{-1}(t)), \qquad t \in Z_m.$$

Let Δ denote a system of distinct representatives for the orbits in Γ_n^m induced by G, and let $\bar{\Delta}$ denote the set of all those elements $\gamma \in \Delta$ for which the character χ is identically 1 on the stabilizer subgroup $G_\gamma = \{\sigma \in G | \sigma(\gamma) = \gamma\}$. Let $\nu(\gamma) = |G_\gamma|$. It is a routine exercise to verify that if $\{e_1, \ldots, e_n\}$ is a basis of V, then the decomposable elements $e_\gamma^* = e_{\gamma(1)} * \ldots * e_{\gamma(m)}$, $\gamma \in \bar{\Delta}$, form a basis of P. In fact, if $\{e_1, \ldots, e_n\}$ is an orthonormal (hereafter abbreviated o.n.) basis of V, then the $|\bar{\Delta}|$ decomposable elements $(|G|/\nu(\gamma)^{\frac{1}{2}})e_\gamma^*$, $\gamma \in \bar{\Delta}$, form an o.n. basis for P with respect to the induced inner product in $\otimes_1^m V$ defined by

$$(x_1 \otimes \ldots \otimes x_m, y_1 \otimes \ldots \otimes y_m) = \prod_{i=1}^{m} (x_i, y_i).$$

In general, if $x_i = \sum_{j=1}^{n} c_{ij} e_j$, $i = 1, \ldots$, m, then the decomposable element $x_1 * \ldots * x_m$ can be expressed in terms of the basis $\{e_\gamma^*, \gamma \in \bar{\Delta}\}$. Given the group G and character χ, we define the generalized matrix function [**3**], d_χ^G, as a mapping from the set of m-square matrices to **C**, by

$$d_\chi^G(B) = \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^{m} b_{i\sigma(i)}. \tag{4}$$

For example, if $G = S_m$ and $\chi = \epsilon$, then $d_\chi^G = \det$; if $G = S_m$ and $\chi \equiv 1$, then $d_\chi^G = \text{per}$. It is a routine calculation to verify that

$$x_1 * \ldots * x_m = \sum_{\gamma \in \bar{\Delta}} \frac{1}{\nu(\gamma)} d_\chi^G(C[1, \ldots, m|\gamma]) e_\gamma^*, \tag{5}$$

where C is the $m \times n$ matrix whose (i, j) entry is c_{ij} and $C[1, \ldots, m|\gamma]$ is the m-square matrix whose (i, j) entry is $c_{i,\gamma(j)}$.

It is an easy task to verify that for arbitrary vectors $x_1, \ldots, x_m, y_1, \ldots, y_m$ in V,

$$(x_1 * \ldots * x_m, y_1 * \ldots * y_m) = \frac{1}{|G|} d_\chi^G([(x_i, y_j)]). \tag{6}$$

If $T \in \mathrm{Hom}(V, V)$, then

$$(7) \qquad \varphi: (v_1, \ldots, v_m) \to Tv_1 * \ldots * Tv_m$$

from $\times_1^m V$ to P is symmetric with respect to G and χ and hence, there is a unique linear map h from P to $U = P$ (see diagram (3)) such that $\varphi = h\mu$. For each linear operator T on V we denote the corresponding linear map h by $K(T)$. Thus for each decomposable element $x_1 * \ldots * x_m$ in P

$$(8) \qquad K(T)x_1 * \ldots * x_m = Tx_1 * \ldots * Tx_m.$$

From (8) we immediately verify for arbitrary S and T in $\mathrm{Hom}(V, V)$ that

$$(9) \qquad K(ST) = K(S)K(T)$$

and

$$(10) \qquad (K(T))^* = K(T^*).$$

If we specialize V to be complex n-tuple space and consider each $n \times n$ complex matrix A to be a linear operator on V, $v \to vA$ for $v \in V$, then with each matrix A we can associate a $|\bar{\Delta}| \times |\bar{\Delta}|$ matrix $K(A)$ defined by (8): if we use the lexicographic ordering in the sequence set $\bar{\Delta}$, and the elements of $\bar{\Delta}$ index the rows and columns of $K(A)$, then the τ, ω entry of the matrix of $K(A)$ relative to the orthonormal basis $\{(|G|/\nu(\gamma))^{\frac{1}{2}} e_\gamma^* | \gamma \in \bar{\Delta}\}$ described above is

$$(11) \qquad (d_\chi{}^G(A[\tau|\omega]))/(\nu(\omega)\nu(\tau))^{\frac{1}{2}}$$

where $B[\tau|\omega]$ means the submatrix of B lying in rows numbered $\tau(1), \ldots, \tau(m)$ and in columns numbered $\omega(1), \ldots, \omega(m)$ [**4**].

Finally, for $v_1, \ldots, v_s$ in V, let $\langle v_1, \ldots, v_s\rangle$ denote the subspace of V spanned by $v_1, \ldots, v_s$.

3. Proofs. In order to prove Theorem 1 we need the following lemma.

LEMMA 1. *If $x_1 * \ldots * x_m = y_1 * \ldots * y_m$, $m < n$, and if $\{x_1, \ldots, x_m\}$ is a linearly independent set, then the sets $\{x_1, \ldots, x_m\}$ and $\{y_1, \ldots, y_m\}$ span the same subspace.*

Proof. Let k be an integer, $1 \leqq k \leqq m$. Since $m < n$, there is a vector $z_k \neq 0$, such that $(y_t, z_k) = 0$ for $t = 1, \ldots, m$. Now let z_i for $i = 1, \ldots, m$, $i \neq k$, be arbitrary vectors in V. Then from (6) we have

$$(12) \qquad (y_1 * \ldots * y_m, z_1 * \ldots * z_m) = \frac{1}{|G|} d_\chi^G([(y_i, z_j)]).$$

Observe that the kth column of the matrix $[(y_i, z_j)]$ is 0 and so the left-hand side of (12) is 0. Thus $(x_1 * \ldots * x_m, z_1 * \ldots * z_m) = 0$ and so $d_\chi{}^G([(x_i, z_j)]) = 0$.

Now choose $z, z_1, \ldots, z_{k-1}, z_{k+1}, \ldots, z_m$ to be biorthogonal to the set $x_k, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_m$. Then the matrix $[(x_i, z_j)]$ has the following form: its first $k - 1$ columns are those of the $m \times m$ identity matrix, its last $m - k$ columns are all zeros, and its kth column consists of the numbers, in order, $(x_1, z_k), \ldots, (x_m, z_k)$. Thus $0 = d_\chi{}^G([(x_i, z_j)]) = (x_k, z_k) = 0$.

https://doi.org/10.4153/CJM-1972-064-8 Published online by Cambridge University Press

Hence we have proved that every vector which is perpendicular to the space spanned by $\{y_i, \ldots, y_m\}$ is perpendicular to the space spanned by $\{x_1, \ldots, x_m\}$. Since $\{x_1, \ldots, x_m\}$ is a linearly independent set, it follows that $\{x_1, \ldots, x_m\}$ and $\{y_1, \ldots, y_m\}$ span the same subspace of V.

Proof of Theorem 1. Clearly if $T = cS$ with $c^m = 1$, then $K(T) = K(S)$. Conversely, assume that $K(T) = K(S)$. Let $\{e_1, \ldots, e_n\}$ be a basis for V such that $\{Te_1, \ldots, Te_r\}$ is a basis for $\operatorname{Im} T$ and $\{e_{r+1}, \ldots, e_n\}$ is a basis for the kernel of T.

Let $x_i = Te_i$ and $y_i = Se_i$ for $i = 1, \ldots, n$ and observe that if $\omega \in \Gamma_n^m$, then

$$(13) \qquad \begin{aligned} x_{\omega(1)} * \ldots * x_{\omega(m)} &= Te_{\omega(1)} * \ldots * Te_{\omega(m)} \\ &= K(T)e_{\omega(1)} * \ldots * e_{\omega(m)} \\ &= K(S)e_{\omega(1)} * \ldots * e_{\omega(m)} \\ &= Se_{\omega(1)} * \ldots * Se_{\omega(m)} \\ &= y_{\omega(1)} * \ldots * y_{\omega(m)}. \end{aligned}$$

For $t = 1, \ldots, m + 1$ let ω^t denote the sequence

$$(1, 2, \ldots, t - 1, t + 1, \ldots, m + 1) \in \Gamma_n^m.$$

Since $m < r = \operatorname{rank} T$ it follows that $\{x_1, \ldots, x_{m+1}\}$ is a linearly independent set, and so we can apply Lemma 1 to (13) to conclude that

$$W_t = \langle x_{\omega^t(1)}, \ldots, x_{\omega^t(m)} \rangle = \langle y_{\omega^t(m)}, \ldots, y_{\omega^t(m)} \rangle, \; t = 1, \ldots, m + 1.$$

Now for each k, $1 \leqq k \leqq r$,

$$\bigcap_{\substack{t=1 \\ t \neq k}}^{m+1} W_t = \langle x_k \rangle = \langle y_k \rangle.$$

Thus Te_k and Se_k span the same space for $k = 1, \ldots, r$. Hence $\{Se_1, \ldots, Se_r\}$ is a linearly independent set and

$$x_j = Te_j = c_j Se_j = c_j y_j$$

for $c_j \neq 0$, $j = 1, \ldots, r$. Therefore

$$\begin{aligned} x_{\omega^t}^* &= y_{\omega^t}^* \\ &= \left(\prod_{\substack{j=1 \\ j \neq t}}^{m+1} c_j \right) x_{\omega^t}^* \\ &\neq 0 \end{aligned}$$

and so

$$\prod_{\substack{j=1 \\ j \neq t}}^{m+1} c_j = 1,$$

for $t = 1, \ldots, m + 1$. Thus $c_1 = \ldots = c_{m+1} = c$ with $c^m = 1$. Similarly we can show that $c_{m+1} = \ldots = c_r = c$. Thus $Te_j = cSe_j$ for $j = 1, \ldots, r$. Now let $k \geqq r + 1$. Then, since $x_k = 0$, we have

$$0 = x_1 * \ldots * x_{t-1} * x_{t+1} * \ldots * x_m * x_k$$
$$= y_1 * \ldots * y_{t+1} * y_{t+1} * \ldots * y_m * y_k.$$

Since the latter tensor is zero, the vectors $y_1, \ldots, y_{t-1}, y_{t+1}, \ldots, y_m, y_k$ must be linearly dependent and so y_k belongs to the intersection of the subspaces spanned by $\{y_1, \ldots, y_{t-1}, y_{t+1}, \ldots, y_m\}$, $t = 1, \ldots, m$. But this intersection is the zero vector and so $y_k = 0$ for $k = r + 1, \ldots, n$. Thus $0 = Te_k = cSe_k$ for $k = r + 1, \ldots, n$, and $T = cS$ with $c^m = 1$.

Proof of Theorem 2. It is easily checked from (9) and (10) that if T is normal then so is $K(T)$. Suppose that rank $T > m$ and $K(T)$ is normal. Then $K(TT^*) = K(T^*T)$ and so by Theorem 1, $TT^* = cT^*T$ for some c with $c^m = 1$. But both TT^* and T^*T are positive semi-definite hermitian operators with the same positive trace. Thus $c = 1$ and so T is normal.

In order to prove Theorem 3 we need two lemmas:

LEMMA 2. *If A is an $n \times n$ matrix of the form*

$$\begin{bmatrix} T & L \\ O & C \end{bmatrix}$$

where T is a $p \times p$ upper triangular matrix and C is an $(n - p) \times (n - p)$ upper triangular matrix with zeros along its main diagonal, then for any $\omega \in \Gamma_n{}^m$ for which $\omega(k) > p$ for some k, $1 \leqq k \leqq m$, the matrix $A[\omega|\omega]$ has a zero row and hence $d_\chi{}^G(A[\omega|\omega]) = 0$.

Proof. Assume that A and ω are as in the statement of the lemma and assume that $\omega(k)$ is the largest of the integers $\omega(1), \ldots, \omega(m)$. Then $\omega(k) > p$. Now the entries in row k of $A[\omega|\omega]$ are in succession $a_{\omega(k)\omega(1)}, \ldots, a_{\omega(k)\omega(m)}$. Since A is upper triangular it follows that $a_{\omega(k)\omega(l)} = 0$ when $\omega(k) > \omega(l)$, and since $\omega(k) > p$ it follows that $a_{\omega(k)\omega(l)} = 0$ when $\omega(l) = \omega(k)$. Since $\omega(k) \geqq \omega(l)$ for $l = 1, \ldots, m$, it follows that all the entries of the kth row of $A[\omega|\omega]$ are zero.

LEMMA 3. *If T is a linear operator on V of rank r, then* rank $K(T) = |\bar{\Delta} \cap \Gamma_r{}^m|$.

Proof. Let $\{u_1, \ldots, u_n\}$ be a basis for V for which $\{Tu_1, \ldots, Tu_r\}$ is a basis for Im T, and $\{u_{r+1}, \ldots, u_n\}$ is a basis for the kernel of T. Let $v_i = Tu_i$ $(i = 1, \ldots, r)$. Then since $B = \{K(T)u_\omega{}^* | \omega \in \bar{\Delta} \cap \Gamma_r{}^m\}$ is a subset of a basis, it is a linearly independent set in Im $K(T)$.

On the other hand if for any k, $1 \leqq k \leqq m$, $\omega(k) > r$, then $K(T)u_\omega{}^* = 0$. Thus B is a basis for Im $K(T)$ and so rank $K(T) = |\bar{\Delta} \cap \Gamma_r{}^m|$.

Proof of Theorem 3. If U is a unitary matrix, then $K(U)$ is unitary and so $K(A)$ is normal if and only if $K(U)^*K(A)K(U) = K(U^*AU)$ is normal. Thus we can assume by Schur's triangularization theorem that A is already an upper triangular matrix of the form described in Lemma 2. Thus $K(A)$ is upper triangular and since we assumed that $K(A)$ is also normal, it follows that $K(A)$ is diagonal. By (11) the main diagonal elements of $K(A)$ are

$$\frac{1}{\nu(\omega)} d_\chi^G(A[\omega|\omega]),\ \omega \in \bar{\Delta}.$$

From Lemma 2 if follows that for any $\omega \in \bar{\Delta}$ for which $\omega(k) > p$ for some k, $1 < k < m$, the corresponding main diagonal element of $K(A)$ is zero. Now rank $A = m$ impliesthat $p \leqq m$. If $p < m$, then by the preceding remarks, rank $K(A) \leqq |\bar{\Delta} \cap \Gamma_p{}^m|$. But if $p < m$, then $\bar{\Delta} \cap \Gamma_p{}^m$ is a proper subset of $\bar{\Delta} \cap \Gamma_m{}^m$ (i.e., $\tau = (1, \ldots, m) \in \bar{\Delta} \cap \Gamma_m{}^m$ but is not in $\bar{\Delta} \cap \Gamma_p{}^m$) and so the rank of $K(A)$ would be less than $|\bar{\Delta} \cap \Gamma_m{}^m|$, contradicting Lemma 3. Thus $p = m$ and it follows that C is the zero matrix (otherwise, rank $A > m$).

We can now assume that A has the form

$$\begin{bmatrix} T & L \\ 0 & 0 \end{bmatrix},$$

where T is $m \times m$ upper triangular with the non-zero eigenvalues $\lambda_1, \ldots, \lambda_m$ of A on the main diagonal. The main diagonal element of $K(A)$ in the position corresponding to the sequence τ is $d_\chi{}^G(A[\tau|\tau]) = \lambda_1 \ldots \lambda_m = \det T$. Now, since K is a representation, the main diagonal element of $K(A)K(A)^*$ corresponding to the sequence τ is

$$\begin{aligned} |\det T|^2 &= (K(A)K(A)^*)_{\tau\tau} \\ &= K(AA^*)_{\tau\tau} \\ &= d_\chi{}^G(AA^*[\tau|\tau]) \\ &\geqq \det(AA^*[\tau|\tau]) \qquad (14) \\ &= \det(TT^* + LL^*) \\ &\geqq \det TT^* \qquad (15) \\ &= |\det T|^2. \end{aligned}$$

The inequality (14) is an application of a result of Schur [**17**] and the inequality (15) is an instance of the result which states that if C and D are positive semi-definite hermitian matrices, then

$$\det(C + D) \geqq \det C. \qquad (16)$$

Moreover if C is positive definite, then it is easy to verify that equality holds in (16) if and only if $D = 0$. Thus from the above we conclude that $LL^* = 0$ and hence that $L = 0$.

4. Applications. In order to derive the Kress–de Vries–Wegmann result, we introduce some additional notation.

For an $n \times n$ complex matrix X, $\lambda_1(X), \ldots, \lambda_n(X)$ will denote the eigenvalues of X, $\alpha_1(X), \ldots, \alpha_n(X)$ will denote the singular values of X, $\lambda(X)$ will denote the n-tuple of eigenvalues of X, and $\alpha(X)$ will denote the n-tuple of singular values of X; if f is a symmetric function on the complex numbers, then $f(\lambda(X))$ will denote

$$(f(\lambda_1(X)), \ldots, f(\lambda_n(X)))$$

and $f(\alpha(X))$ will denote

$$(f(\alpha_1(X)), \ldots, f(\alpha_n(X)));$$

$E_r(t_1, \ldots, t_n)$ will denote the rth elementary symmetric function of $t_1, \ldots, t_n$ and $C_m(X)$ will denote the mth compound matrix of X, i.e., $C_m(X)$ is just $K(X)$ with G the full symmetric group of degree m and χ the alternating character on G. The eigenvalues of $C_m(A)$ are just the numbers

$$\lambda_\omega(C_m(A)) = \prod_{i=1}^{m} \lambda_{\omega(i)}(A)$$

as ω runs over $Q_{m,n}$, the set of strictly increasing sequences of length m of integers chosen from $1, \ldots, n$. Thus trace $(C_m(A))$ is just $E_m(\lambda(A))$.

LEMMA 4.

$$(17) \quad ||A||^4 - \tfrac{1}{2}||D||^2 - \left(\sum_{i=1}^{n} |\lambda_i(A)|^2\right)^2 = E_1((\alpha(A^2))^2) - E_1(|\lambda(A^2)|^2) + 2(E_2((\alpha(A))^2) - E_2(|\lambda(A)|^2).$$

Proof. We compute

$$||A||^4 = \left(\sum_{i=1}^{n} \alpha_i^2(A)\right)^2$$
$$= \sum_{i=1}^{n} \alpha_i^4(A) + 2E_2((\alpha(A))^2);$$
$$||D||^2 = \text{trace } ((AA^* - A^*A)^2)$$
$$= \text{trace } ((AA^*)^2 + (A^*A)^2 - AA^*A^*A - A^*AAA^*)$$
$$= 2 \text{ trace } ((AA^*)^2) - 2 \text{ trace } (A^2A^{*2})$$
$$= 2\sum_{i=1}^{n} \lambda_i^2(AA^*) - 2\sum_{i=1}^{n} \alpha_i^2(A^2)$$
$$= 2\sum_{i=1}^{n} \alpha_i^4(A) - 2\sum_{i=1}^{n} \alpha_i^2(A^2);$$
$$\left(\sum_{i=1}^{n} |\lambda_i(A)|^2\right)^2 = \sum_{i=1}^{n} |\lambda_i(A)|^4 + 2E_2(|\lambda(A)|^2)$$
$$= E_1(|\lambda(A^2)|^2) + 2E_2(|\lambda(A)|^2).$$

Thus the left side of (17) is equal to

$$(18) \quad E_1((\alpha(A^2))^2) - E_1(|\lambda(A^2)|^2) + 2[E_2((\alpha(A))^2) - E_2(|\lambda(A)|^2)].$$

We obtain the inequality (1) by rewriting (18) to obtain

$$(19)\quad \left[\|A^2\|^2 - \sum_{i=1}^{n} |\lambda_i(A^2)|^2\right] + 2\left[\|C_2(A)\|^2 - \sum_{\omega \in Q_{2,n}} |\lambda_\omega(C_2(A))|^2\right]$$

and applying Schur's inequality to both A^2 and $C_2(A)$.

Equality holds in (1) if and only if (19) is 0. But from Schur's inequality (19) is 0 if and only if both A^2 and $C_2(A)$ are normal.

Suppose equality holds in (1) and suppose A is not normal. Then by Theorem 2, rank $A \leqq 2$. If rank $A = 2$, then by Theorem 3, A is unitarily similar to a matrix of the form

$$B = \begin{bmatrix} \lambda_1 & a \\ 0 & \lambda_2 \end{bmatrix} \oplus 0_{n-2}$$

where $a \neq 0$ because A is not normal. Since A^2 is normal so is B^2 and so we conclude that $\lambda_2 = -\lambda_1$. Thus A is unitarily similar to a matrix $L \oplus 0_{n-2}$ where

$$L = \begin{bmatrix} \lambda & a \\ 0 & -\lambda \end{bmatrix},$$

with $\lambda \neq 0$. We wish to show that $L \oplus 0_{n-2}$ is unitarily similar to a matrix of the form in (2) and conversely that any matrix of the form in (2) is unitarily similar to a matrix of the form $L \oplus 0_{n-2}$. The converse is just a consequence of Schur's triangularization theorem and the fact that we are assuming that A is not normal.

Now $\{\text{trace}(X), \text{trace}(X^2), \text{trace}(X^*X)\}$ is a complete set of unitary invariants for complex 2×2 matrices X (see [**6**]). Let S denote the matrix in (2). Then using the fact that v and w are orthonormal column vectors it is a routine matter to calculate that $\text{trace}\, S = 0$, $\text{trace}(S^2) = 2r\alpha^2$, and $\text{trace}(S^*S) = |\alpha|^2(1 + r^2)$. Now $\text{trace}\, L = 0$, $\text{trace}\, L^2 = 2\lambda^2$, and $\text{trace}\, L^*L = 2|\lambda|^2 + |a|^2$. Thus the problem of showing that a matrix of the form $L \oplus 0_{n-2}$ is unitarily similar to a matrix of the form S consists of the following: given non-zero complex numbers λ and a, is there a non-zero complex number α and a real number r, $0 < r < 1$, such that

$$r\alpha^2 = \lambda^2,\ 2|\lambda|^2 + |a|^2 = |\alpha|^2(1 + r^2)?$$

But it is a routine calculation to check that the answer to this last question is "yes".

If rank $A = 1$, then $A = \alpha xy^*$ for some complex number $\alpha \neq 0$ and for some complex column n-tuples x and y with $1 = \|x\|^2 = x^*x = y^*y = \|y\|^2$. Since, by assumption, A is not normal it follows that

$$(20)\quad |\alpha|^2 xx^* = AA^* \neq A^*A = |\alpha|^2 yy^*.$$

But A^2 is normal and so

$$|\alpha|^4 (x, y)(y, x) xx^* = A^2(A^2)^* = (A^2)^* A^* = |\alpha|^4 (x, y)(y, x) yy^*.$$

Hence we conclude from (20) that $(x, y) = 0$ and so A has the form (2) with $r = 0$.

References

1. Rainer Kress, Hans Ludwig de Vries, and Rudolf Wegmann, *On non-normal matrices* (to appear in Linear Algebra and Appl.).
2. Marvin Marcus and Henryk Minc, *Introduction to linear algebra* (Macmillan Company, New York, 1965).
3. ——— *Generalized matrix functions*, Trans. Amer. Math. Soc. *116* (1965), 316–329.
4. ——— *Permutations on symmetry classes*, J. Algebra *5* (1967), 59–71.
5. Marvin Marcus, *On two classical results of I. Schur*, Bull. Amer. Math. Soc. *70* (1964), 685–688.
6. F. D. Murnaghan, *On the unitary invariants of a square matrix*, An. Acad. Brasil. Ci. *26* (1954), 1–7.

University of California,
Santa Barbara, California;
University of Victoria,
Victoria, British Columbia

https://doi.org/10.4153/CJM-1972-064-8 Published online by Cambridge University Press