Conditions for the Generalized Numerical Range to Be Real*

Marvin Marcus

Department of Mathematics and Department of Computer Science
Institute for the Interdisciplinary Applications
of Algebra and Combinatorics
University of California at Santa Barbara
Santa Barbara, California 93106

and

Markus Sandy

Department of Mathematics
Institute for the Interdisciplinary Applications
of Algebra and Combinatorics
University of California at Santa Barbara
Santa Barbara, California 93106

Dedicated to Helmut Wielandt on his 75th birthday

Submitted by Emeric Deutsch

ABSTRACT

Necessary and sufficient conditions are given for the C-numerical range of a matrix A to be a subset of the real axis. In particular, it is shown that both A and C must be translates of hermitian matrices.

1. INTRODUCTION

Let C and A be n-square complex matrices, and let $c = a + ib$ be a complex n-tuple. The c-numerical range of A [1] is the subset of the complex plane

$$W_c(A) = \left\{ \sum_{k=1}^{n} c_k(Ax_k, x_k) \,\middle|\, x_1, \dots, x_n \text{ o.n.} \right\}. \tag{1}$$

*The research of both authors was supported by the Air Force Office of Scientific Research under grant #AFOSR-83-0150.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/85/$3.30

In (1) the standard inner product is used and $x_1, \ldots, x_n$ run over all sets of orthonormal (o.n.) bases. The C-numerical range of A [1] is the set of numbers

$$W(C, A) = \{ \operatorname{tr}(CU^*AU) | U \text{ unitary} \}. \tag{2}$$

In (2) U runs over all n-square unitary matrices.

There are a number of elementary facts about $W_c(A)$ and $W(C, A)$ that we require. First note that if $c = e_k$, where e_k is the n-tuple whose kth component is 1 with the remaining components 0, then $W_c(A)$ is the classical Toeplitz-Hausdorff numerical range (field of values) $W(A)$ [12, 4, 11, 6, 7, 3]. If U and V are arbitrary unitary matrices, then

$$W_c(U^*AU) = W_c(A), \tag{3}$$

and

$$W(U^*CU, V^*AV) = W(C, A). \tag{4}$$

Also

$$W(C, A) = W(A, C), \tag{5}$$

and if C is normal with eigenvalues $c_1, \ldots, c_n$, then

$$W(C, A) = W_c(A). \tag{6}$$

If $A = H + iK$ is the hermitian decomposition of A, then the reality of $W(A)$, i.e., the condition

$$W(A) \subset \mathbb{R}, \tag{7}$$

implies that $K = 0$ and hence that A is hermitian. The converse is equally obvious. The purpose of this paper is to obtain necessary and sufficient conditions on c, A, C in order that

$$W_c(A) \subset \mathbb{R}, \tag{8}$$

and

$$W(C, A) \subset \mathbb{R}. \tag{9}$$

Interest in the geometric structure of the numerical range and its generalizations dates to the original Toeplitz-Hausdorff theorem, in which $W(A)$ is proved convex. Later $W_c(A)$ was proved convex for real c [13], and the relation between algebraic properties of A and geometric properties of the generalized numerical ranges of A was studied. An example of this kind of result states that if $W(A)$ is a subset of the unit disk and every eigenvalue of A has modulus 1 then A is unitary [9]. Another result of this type shows that although $W(A)$ is the convex polygon spanned by the eigenvalues of A when A is normal, the converse is true only for $n \leqslant 4$ [5]. However, if several of the higher numerical ranges of A are convex polygons, then A is normal [8]. In [2] it is proved that if

$$r_C(A) = \max_{z \in W(C,A)} |z|,$$

then if C is not a scalar matrix and $\operatorname{tr}(C) \neq 0$ then $r_C(A) = 0$ implies $A = 0$. In other words, if $W(C, A)$ is the origin, then $A = 0$ under these conditions. Various proofs of this result subsequently appeared in [10].

Suppose that $c = \gamma e$, where γ is a complex number and e is the n-tuple with every component equal to 1. Then

$$W_c(A) = \gamma \operatorname{tr}(A),$$

a single point. Similarly, if C is a scalar matrix, $C = \gamma I_n$, then

$$W(C, A) = \gamma \operatorname{tr}(A)$$

and either of the conditions (8) and (9) collapses to $\gamma \operatorname{tr}(A) \in \mathbb{R}$. Of course, if A is scalar, then $W_c(A)$ and $W(C, A)$ again collapse to a single point. *Henceforth we shall assume that c is not a multiple of e and that neither A nor C is a scalar matrix.*

2. STATEMENTS

In the present section we state the main results of this paper, and in Section 3 the supporting lemmas and proofs appear.

THEOREM 1. *If $W(C, A)$ is contained in a fixed line parallel to the real axis, then at least one of C and A must be normal.*

THEOREM 2. *The inclusion*

$$W_c(A) \subset \mathbb{R} \tag{10}$$

holds iff one of the following two independent sets of conditions is satisfied:

I. *There exist μ, τ, ρ in $\mathbb{R}$ and a hermitian matrix L such that*

$$b \text{ and } e \text{ are linearly independent}, \tag{11}$$

$$a + \mu b = \rho e, \tag{12}$$

$$A = \tau I_n + (\mu + i)L, \tag{13}$$

$$\tau \sum_{k=1}^{n} b_k + \rho \operatorname{tr}(L) = 0. \tag{14}$$

II. *There exist β and ρ in $\mathbb{R}$ and a hermitian matrix L such that*

$$b = \beta e, \tag{15}$$

$$A = i\rho I_n + L, \tag{16}$$

$$\rho \sum_{k=1}^{n} a_k + \beta \operatorname{tr}(L) = 0. \tag{17}$$

THEOREM 3. *The inclusion*

$$W(C, A) \subset \mathbb{R} \tag{18}$$

holds iff both C and A are normal and one of the two independent sets of conditions is satisfied:

I′. *There exist μ, τ, ρ in $\mathbb{R}$ and hermitian matrices L and T such that*

$$T \text{ is not scalar}, \tag{19}$$

$$C = \rho I_n + (i - \mu)T, \tag{20}$$

$$A = \tau I_n + (\mu + i)L, \tag{21}$$

$$\operatorname{tr}(\tau T + \rho L) = 0. \tag{22}$$

II′. *There exist β and ρ in $\mathbb{R}$ and hermitian matrices L and S such that*

$$C = i\beta I_n + S, \tag{23}$$

$$A = i\rho I_n + L, \tag{24}$$

$$\operatorname{tr}(\beta L + \rho S) = 0. \tag{25}$$

Note that in I′ the imaginary part of C is not scalar, whereas in II′ it is. Recall from Section 1 that we have assumed neither C nor A is scalar, and hence neither L nor S in I′ and II′ above can be scalar.

For the purpose of the final result we drop the conditions that neither C nor A is scalar.

THEOREM 4. *The classical numerical range satisfies*

$$W(CU^*AU) \subset \mathbb{R} \tag{26}$$

for all unitary U iff at least one of C and A is scalar and their product is hermitian.

3. PROOFS

We begin with two lemmas necessary for the proof of Theorem 1.

LEMMA 1. *If A is not normal, then there exists a unitary matrix R such that R^*AR is upper-triangular (u.t.) and the 1,2 entry satisfies $(R^*AR)_{12} \neq 0$.*

Proof. If A is 2-square, then the conclusion is clear, since the 1,2 entry in any u.t. unitary transform of A is not 0. Henceforth, assume A is n-square, $n \geqslant 3$. By Schur's [7] theorem bring A to u.t. form, and if there are any multiplicities among the eigenvalues $\lambda_1, \dots, \lambda_n$ of A, assume that they occur among $\lambda_1, \dots, \lambda_{n-1}$, i.e., assume that if any multiplicities occur, at least two of the eigenvalues $\lambda_1, \dots, \lambda_{n-1}$ are equal. If the upper left $(n-1)$-square principal submatrix in A is not diagonal, then it is not normal, and a simple induction on n completes the proof. Thus we may assume A has the form

$$A = \left[\begin{array}{cccccc|c} \lambda_1 & & & & & & a_1 \\ & \lambda_2 & & & 0 & & a_2 \\ & & \ddots & & & & \vdots \\ & & & \lambda_p & & & a_p \\ & 0 & & & \ddots & & \vdots \\ & & & & & \lambda_{n-1} & a_{n-1} \\ \hline & & & 0 & & & \lambda_n \end{array}\right]. \tag{27}$$

If some a_p, $p \leqslant n-1$, in (27) were 0, then interchange rows 1 and p and columns 1 and p, a unitary similarity on A, so that we may assume under these circumstances that

$$A = \left[\begin{array}{cccc|c} \lambda_1 & & & & 0 \\ & \lambda_2 & 0 & & a_2 \\ & 0 & \ddots & & \vdots \\ & & & \lambda_{n-1} & a_{n-1} \\ \hline & & 0 & & \lambda_n \end{array}\right]. \tag{28}$$

The lower right $(n-1)$-square principal submatrix in (28) is not normal (since A is not normal), and hence by induction on n this submatrix is unitarily similar to an u.t. matrix with 1,2 entry not 0. We could then perform a block permutation to place this submatrix in the upper left position with λ_1 in the n,n position, to complete the proof. Thus we can assume in (27) that none of $a_1,\ldots,a_{n-1}$ is 0. Moreover, if there are any multiplicities among $\lambda_1,\ldots,\lambda_{n-1}$, we can also assume that $\lambda_1 = \lambda_2 = \lambda$, so that A has the form

$$A = \left[\begin{array}{cc|ccc|c} \lambda & & & 0 & & a_1 \\ & \lambda & & & & a_2 \\ \hline & & \lambda_3 & & 0 & a_3 \\ & 0 & & \ddots & & \vdots \\ & & 0 & & \lambda_{n-1} & a_{n-1} \\ \hline & 0 & & 0 & & \lambda_n \end{array}\right]. \tag{29}$$

Since a_1 and a_2 are not 0, choose a unitary 2-square matrix W such that

$$W^*\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ a \end{bmatrix}, \qquad a \neq 0. \tag{30}$$

Let $U = W \dot{+} I_{n-2}$, and we compute from (29) and (30) that U^*AU has the same form as (28) and the proof can be completed as before.

Thus we can assume that A has the form (27) with $\lambda_1, \ldots, \lambda_n$ distinct and $a_1, \ldots, a_{n-1}$ different from 0. We prove that any such matrix cannot have an eigenvector corresponding to λ_n whose nth component is 0. For, suppose that

$$Ax = \lambda_n x \tag{31}$$

and $x_n = 0$. Then componentwise the equation (31) becomes

$$\lambda_t x_t = \lambda_n x_t, \qquad t = 1, \ldots, n-1. \tag{32}$$

But $\lambda_t \neq \lambda_n$ and hence $x_t = 0$, $t = 1, \ldots, n-1$, a contradiction. Let T denote the lower right principal $(n-1)$-square submatrix of A in (27). Since T has precisely the same form as A, we may apply the preceding remark to conclude that any eigenvector of T corresponding to λ_n must have its last coordinate z not 0. Choose a unitary $(n-1)$-square W so that W^*TW is u.t. with λ_n in the 1,1 position. Then the first column of W is an eigenvector of T corresponing to λ_n. Let $V = 1 \dot{+} W$, and compute that

$$V^*AV = \left[\begin{array}{c|c} \lambda_1 & [0 \ \cdots \ 0 \ a_1]W \\ \hline 0 & W^*TW \end{array}\right]. \tag{33}$$

Note that

$$[0 \quad \cdots \quad 0 \quad a_1]W = [a_1 z \quad * \quad \cdots \quad *],$$

in which $a_1 z \neq 0$. This completes the proof. ∎

LEMMA 2. *For $n = 2$, if $W(C, A)$ is contained in a fixed line parallel to the real axis, then at least one of C and A must be normal.*

Proof. From (4) we may assume that

$$C=\begin{bmatrix} c_{11} & c_{12} \\ 0 & c_{22} \end{bmatrix}, \qquad A=\begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}.$$

Now let

$$U=\begin{bmatrix} 0 & e^{i\phi} \\ -e^{-i\phi} & 0 \end{bmatrix},$$

and compute that

$$CU^*AU=\begin{bmatrix} c_{11}a_{22}-c_{12}a_{12}e^{-i2\phi} & c_{12}a_{11} \\ -c_{22}a_{12}e^{-i2\phi} & c_{22}a_{11} \end{bmatrix}.$$

Thus

$$\operatorname{tr}(CU^*AU)=-c_{12}a_{12}e^{-i2\phi}+(c_{11}a_{22}+c_{22}a_{11}). \tag{34}$$

If neither C nor A is normal, $c_{12}a_{12}\neq 0$, and since ϕ is arbitrary, the numbers (34) do not lie on a line parallel to the real axis. ■

Proof of Theorem 1. If we assume that neither C nor A is normal, then by Lemma 1 we may assume that

$$C=\begin{bmatrix} M & L \\ 0 & N \end{bmatrix}, \qquad A=\begin{bmatrix} Q & F \\ 0 & G \end{bmatrix}, \tag{35}$$

in which M and Q are 2-square with $M_{12}Q_{12}\neq 0$, and both C and A are u.t. Let U be an n-square unitary matrix of the form

$$U=R\dotplus I_{n-2}$$

in which R is 2-square unitary. We compute that

$$CU^*AU=\begin{bmatrix} MR^*QR & MR^*F+LG \\ 0 & NG \end{bmatrix}, \tag{36}$$

and hence

$$\operatorname{tr}(CU^*AU) = \operatorname{tr}(MR^*QR) + \operatorname{tr}(NG) \tag{37}$$

We are assuming that $\operatorname{tr}(CU^*AU)$ lies on a fixed line parallel to the real axis. Since $\operatorname{tr}(NG)$ is a constant, we conclude from (37) that

$$\operatorname{tr}(MR^*QR) \tag{38}$$

lies on a fixed line parallel to the real axis as R varies over all 2-square unitary matrices. But then Lemma 2 implies that at least one of M or Q is normal and hence $M_{12}Q_{12}$ must be 0, a contradiction. ■

LEMMA 3. *If M and N are n-square hermitian then*

$$(Mx_1, x_1) = (Nx_2, x_2) \tag{39}$$

for all o.n. x_1, x_2 iff

Case 1 ($n = 2$):

$$M + N = \lambda I_2 \qquad \textit{where} \quad \lambda = \operatorname{tr} M = \operatorname{tr} N.$$

Case 2 ($n \geqslant 3$):

$$M = N = \lambda I_n, \qquad \lambda \textit{ real}.$$

Proof. *Case 1* ($n = 2$): Replace x_1 and x_2 in (39) by

$$sx_1 + tx_2 \tag{40a}$$

and

$$-\bar{t}x_1 + \bar{s}x_2 \tag{40b}$$

respectively, where s and t are complex numbers satisfying $|s|^2 + |t|^2 = 1$. The equation (39) then becomes

$$(M(sx_1 + tx_2), sx_1 + tx_2) = (N(-\bar{t}x_1 + \bar{s}x_2), -\bar{t}x_1 + \bar{s}x_2),$$

which, using (39), simplifies to

$$\operatorname{Re} s\bar{t}((M+N)x_1, x_2) = 0 \tag{41}$$

for all $|s|^2 + |t|^2 = 1$ and o.n. x_1 and x_2. Since $s\bar{t}$ can have any argument, it follows that (41) holds iff

$$((M+N)x_1, x_2) = 0 \tag{42}$$

for all x_1, x_2 o.n. We now use the following general fact:

$$(Rx, y) = 0$$

for all x and y for which $(x, y) = 0$ iff R is a multiple of I_n. Hence we can conclude

$$M + N = \lambda I_2. \tag{43}$$

For $n = 2$,

$$\begin{aligned} \operatorname{tr}(M) &= (Mx_1, x_1) + (Mx_2, x_2) \\ &= (Nx_2, x_2) + (Nx_1, x_1) \\ &= \operatorname{tr}(N). \end{aligned}$$

Then

$$\operatorname{tr}(M+N) = 2\lambda,$$

$$\operatorname{tr}(M) = \lambda,$$

and similarly

$$\operatorname{tr}(N) = \lambda.$$

Conversely, if $M + N = \lambda I_2$ and $\lambda = \operatorname{tr} M = \operatorname{tr} N$, then

$$\begin{aligned} (Mx_1, x_1) &= \operatorname{tr}(M) - (Mx_2, x_2) = \lambda - (Mx_2, x_2) \\ &= ((\lambda I_2 - M)x_2, x_2) \\ &= (Nx_2, x_2). \end{aligned}$$

Case 2 ($n \geqslant 3$): Let u be an arbitrary unit vector. Then since $n \geqslant 3$, we may choose v and w such that the triple u, v, w are o.n. Then

$$(Nx, x) = (Mv, v) \qquad \text{for all unit vectors } x \in \langle v \rangle^{\perp}. \tag{44}$$

Thus letting $c_v = (Mv, v)$ we can conclude from (44) that

$$PNx = c_v x, \qquad x \in \langle v \rangle^{\perp}, \tag{45}$$

where P is the orthogonal projection into $\langle v \rangle^{\perp}$. Similarly

$$QNx = c_w x, \qquad x \in \langle w \rangle^{\perp}, \tag{46}$$

where Q is the orthogonal projection into $\langle w \rangle^{\perp}$. Now replace x by u in (45) and (46) to obtain

$$PNu = c_v u, \tag{47}$$

$$QNu = c_w u. \tag{48}$$

If we set $y = Nu$, then we can write

$$y = aw + bv + cu + z, \qquad z \in \langle u, v, w \rangle^{\perp}.$$

Then from (47),

$$c_v u = Py = aw + cu + z$$

and hence $a = 0$ and $z = 0$, so that

$$y = bv + cu.$$

From (48)

$$c_w u = Qy = bv + cu,$$

and hence $b = 0$. But then $y = cu$, and thus

$$Nu \in \langle u \rangle.$$

But u was arbitrary, so that N must be a multiple of the identity, say $N = \lambda I_n$. Similarly M is a multiple of the identity and

$$M = N = \lambda I_n$$

follows from (39). ■

As a consequence of Lemma 3 we have the following result.

Corollary. *If M is hermitian and $(Mx_1, x_1) = (Mx_2, x_2)$ for all o.n. x_1 and x_2, then $M = \lambda I_n$.*

We are now in a position to prove Theorem 2.

Proof of Theorem 2. We compute that

$$\sum_{k=1}^{n} c_k(Ax_k, x_k) = \sum_{k=1}^{n} (a_k + ib_k)\{(H_k, x_k) + i(Kx_k, x_k)\}$$

$$= \sum_{k=1}^{n} ((a_k H - b_k K)x_k, x_k)$$

$$+ i \sum_{k=1}^{n} ((b_k H + a_k K)x_k, x_k). \tag{49}$$

Hence (49) is real iff

$$\sum_{k=1}^{n} (M_k x_k, x_k) = 0, \tag{50}$$

where $M_k = b_k H + a_k K$, $k = 1, \ldots, n$. If the left-hand side of (50) is evaluated for the ordered sequence of o.n. vectors

$$x_1, x_2, \ldots, x_{i-1}, x_j, x_{i+1}, \ldots, x_{j-1}, x_i, x_{j+1}, \ldots, x_n, \tag{51}$$

then from (50)

$$(M_i x_i, x_i) + (M_j x_j, x_j) = (M_i x_j, x_j) + (M_j x_i, x_i),$$

or

$$((M_i - M_j)x_i, x_i) = ((M_i - M_j)x_j, x_j) \tag{52}$$

in which x_i and x_j are arbitrary o.n. vectors. Applying the Corollary, it follows that

$$M_i - M_j = \lambda_{ij} I_n, \tag{53}$$

where $\lambda_{ij} \in \mathbb{R}$, and (53) holds for all $i < j$. Now

$$M_i - M_j = (b_i - b_j)H - (a_i - a_j)K = \lambda_{ij} I_n,$$

or

$$(b_i - b_j)H = \lambda_{ij} I_n + (a_i - a_j)K. \tag{54}$$

There are two distinct cases to consider:

I. There exists a pair $p < q$ such that $b_p - b_q \neq 0$, i.e., b and e are linearly independent.

II. $b = \beta e$.

Case I: We can write

$$H = \frac{\lambda_{pq}}{b_p - b_q} I_n + \frac{a_p - a_q}{b_p - b_q} K,$$

or

$$H = \tau I_n + \mu K, \tag{55}$$

where $\tau, \mu \in \mathbb{R}$. Replace H in (50) by the expression in (55) to obtain

$$\sum_{k=1}^{n} (M_k x_k, x_k) = \tau \sum_{k=1}^{n} b_k + \sum_{k=1}^{n} (a_k + \mu b_k)(K x_k, x_k) = 0 \tag{56}$$

Evaluating (56) on the n-tuple of o.n. vectors (51) and subtracting the result from (56) yields

$$\{(a_i - a_j) + \mu(b_i - b_j)\}\{(K x_i, x_i) - (K x_j, x_j)\} = 0 \tag{57}$$

for all o.n. pairs of vectors x_i, x_j. If

$$(Kx_i, x_i) - (Kx_j, x_j)$$

were to vanish for all o.n. pairs of vectors x_i, x_j, it would follow that K would be scalar and hence from (55) that H, and finally A, would be scalar. But this case was dealt with before, and we are assuming that A is not scalar. Hence the first factor in (57) is 0 for any pair of integers $i < j$. Thus for any pair $i < j$

$$a_i + \mu b_i = a_j + \mu b_j, \tag{58}$$

or

$$a + \mu b = \rho e. \tag{59}$$

Substituting (59) into (56) results in

$$\tau \sum_{k=1}^{n} b_k + \rho \operatorname{tr}(K) = 0. \tag{60}$$

Also note from (25) that

$$A = H + iK = \tau I_n + (\mu + i)K. \tag{61}$$

The conditions (59), (60), and (61) are precisely (12), (14), and (13) respectively with the generic L replacing the letter K. Conversely we can easily confirm that the conditions I immediately imply that

$$W_c(A) \subset \mathbb{R}.$$

Case II ($b = \beta e$): Since we are assuming that $c \neq \gamma e$, it follows from (54) that for some $\rho \in \mathbb{R}$

$$K = \rho I_n$$

and so

$$A = H + iK = H + i\rho I_n. \tag{62}$$

Thus

$$\begin{aligned}\sum_{k=1}^{n} c_k(Ax_k, x_k) &= \sum_{k=1}^{n} (a_k + i\beta)((H + i\rho I_n)x_k, x_k)\\ &= \sum_{k=1}^{n} a_k(Hx_k, x_k) - n\rho\beta\\ &\quad + i\left\{\beta\,\mathrm{tr}(H) + \rho\sum_{k=1}^{n} a_k\right\}.\end{aligned}$$

From (10) we have that

$$\beta\,\mathrm{tr}(H) + \rho\sum_{k=1}^{n} a_k = 0. \tag{63}$$

Thus if $W_c(A) \subseteq \mathbb{R}$ and $b = \beta e$, then the conditions (62) and (63) hold for appropriate real numbers β and ρ and an hermitian matrix L [H is to be replaced by L in (62) and (63)]. The conditions (62) and (63) are then precisely (16) and (17). Conversely if (15), (16), and (17) hold, then we compute that

$$\begin{aligned}\sum_{k=1}^{n} c_k(Ax_k, x_k) &= \sum_{k=1}^{n} (a_k + i\beta)\{(Lx_k, x_k) + i\rho\}\\ &= \sum_{k=1}^{n} a_k(Lx_k, x_k) - n\beta\rho\\ &\quad + i\left\{\beta\sum_{k=1}^{n} (Lx_k, x_k) + \rho\sum_{k=1}^{n} a_k\right\}.\end{aligned} \tag{64}$$

The imaginary part of (64) is

$$\beta\,\mathrm{tr}(L) + \rho\sum_{k=1}^{n} a_k,$$

which by (17) is 0. ∎

We proceed to the proof of Theorem 3.

Proof of Theorem 3. Recall that we are assuming that neither C nor A is scalar. By Theorem 1 the inclusion (18) implies that at least one of C or A must be normal. Assume initially that C is normal and that the eigenvalues of C are $c_k = a_k + ib_k$, $k = 1, \ldots, n$, i.e.,

$$c = a + ib. \tag{65}$$

Then since $W(C, A)$ is invariant under unitary similarity, choose a unitary V such that

$$V^*CV = \operatorname{diag}(c_1, \ldots, c_n) \tag{66}$$

and observe that

$$\begin{aligned} W(C, A) &= W(V^*CV, A) \\ &= W(\operatorname{diag}(c_1, \ldots, c_n), A) \end{aligned}$$

An element of $W(C, A)$ has the form

$$\begin{aligned} \operatorname{tr}(\operatorname{diag}(c_1, \ldots, c_n) U^*AU) &= \sum_{k=1}^{n} c_k (U^*AU)_{kk} \\ &= \sum_{k=1}^{n} c_k (U^*AUe_k, e_k) \\ &= \sum_{k=1}^{n} c_k (AUe_k, Ue_k) \\ &= \sum_{k=1}^{n} c_k (Ax_k, x_k), \end{aligned} \tag{67}$$

where $x_1, \ldots, x_n$ run over all o.n. sets of n vectors as U runs over all unitary matrices.

Thus

$$W(C, A) = W_c(A),$$

and the condition (18) becomes (10), and we are able to apply Theorem 2 once we now that c is not a multiple of e. But since C is normal and not a scalar, this follows immediately. Hence by Theorem 2 one of the following sets of conditions must hold:

I. There exist real numbers μ, τ, ρ in $\mathbb{R}$ and a hermitian matrix L such that

$$b \text{ and } e \text{ are linearly independent,} \tag{68}$$

$$a + \mu b = \rho e, \tag{69}$$

$$A = \tau I_n + (\mu + i)L, \tag{70}$$

$$\tau \sum_{k=1}^{n} b_k + \rho \operatorname{tr}(L) = 0. \tag{71}$$

II. There exist β and ρ in $\mathbb{R}$ and a hermitian matrix L such that

$$b = \beta e, \tag{72}$$

$$A = i\rho I_n + L, \tag{73}$$

$$\rho \sum_{k=1}^{n} a_k + \beta \operatorname{tr}(L) = 0. \tag{74}$$

Assume first that I holds. Write

$$C = S + iT. \tag{75}$$

Then the eigenvalues of C are c, and the normality implies that the eigenvalues of S are a and those of T are b. The statement (68) then becomes: T is not scalar and

$$S + \mu T = \rho I_n. \tag{76}$$

Thus substituting (76) in (75) results in (19) and (20). Also (70) is precisely (21). Similarly (71) becomes

$$\tau \operatorname{tr}(T) + \rho \operatorname{tr}(L) = 0,$$

the condition (22).

Conversely the conditions (19), (20), (21), and (22) imply that

$$W(A, C) = W_c(A) \subseteq \mathbb{R}. \tag{77}$$

For, a typical element of $W(C, A)$ is

$$\begin{aligned}\operatorname{tr}(CU^*AU) &= \operatorname{tr}([\rho I_n + (\mu - i)T]U^*[\tau I_n + (\mu + i)L]U) \\ &= \rho\tau n + \rho\mu \operatorname{tr}(L) - \mu\tau \operatorname{tr}(T) - (\mu^2 + 1)\operatorname{tr}(TU^*LU) \\ &\quad + i \operatorname{tr}(\tau I_n + \rho L). \end{aligned} \tag{78}$$

Thus (77) holds iff (22) holds.

Assume now that II [i.e., (72), (73), (74)] holds. Then (72) implies that $T = \beta I_n$. The condition (73) is precisely (24), and (74) becomes

$$\operatorname{tr}(\beta L + \rho S) = 0,$$

precisely (25).

We have proved that (18) implies I′ or II′ holds. We have also shown in (78) that I′ implies that $W(C, A) \subset \mathbb{R}$. It remains to show that II′ implies that (18) holds. Again we compute that a typical element of $W(C, A)$ is

$$\begin{aligned}\operatorname{tr}(CU^*AU) &= \operatorname{tr}((i\beta I_n + S)U^*(i\rho I_n + L)U) \\ &= -n\beta\rho + \operatorname{tr}(SU^*LU) \\ &\quad + i \operatorname{tr}(\beta L + \rho S). \end{aligned} \tag{79}$$

The condition (25) in II′ guarantees that (18) holds.

Observe that had we assumed initially that A were normal, the argument could have proceeded exactly as above with the roles of A and C interchanged. ■

Proof of Theorem 4. As was previously mentioned, the numerical range of an n-square matrix is real if and only if that matrix is hermitian and so it follows from (26) that if either A or C is scalar, then their product must be hermitian. We will henceforth assume that neither A nor C is scalar, in order to obtain a contradiction. The hypothesis (26) implies that

$$(CU^*AUx, x) \in \mathbb{R}$$

for all unit vectors x and unitary matrices U, and so

$$\operatorname{tr}(CU^*AU) \in \mathbb{R}$$

for all such U. But then Theorem 3 implies that both A and C are normal matrices. Write $C = VDV^*$, $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, and note that

$$\begin{aligned} W(CU^*AU) &= W(VDV^*U^*AU) \\ &= W(D(UV)^*A(UV)) \\ &= W(DR^*AR) \end{aligned}$$

is always real (here $R = UV$ is unitary). Therefore

$$\begin{aligned} (DR^*ARe_k, e_k) &= (R^*ARe_k, D^*e_k) \\ &= (R^*ARe_k, \bar{\lambda}_k e_k) \\ &= (\lambda_k ARe_k, Re_k) \end{aligned}$$

is always real for all unitary matrices R, and thus $\lambda_k A$ is hermitian, $k = 1, \dots, n$.

Since we are assuming that C is not scalar, we can conclude that for some nonzero eigenvalue λ of C

$$A = \lambda H, \qquad H \text{ hermitian.}$$

A similar argument shows that

$$C = \mu K, \qquad K \text{ hermitian,}$$

for some $\mu \neq 0$. Thus we have that

$$W(CU^*AU) = \lambda\mu W(KU^*HU) \subseteq \mathbb{R},$$

for all unitary U. To abbreviate the notation set $[k] = \operatorname{diag}(k_1, \dots, k_n)$, $[h] = \operatorname{diag}(h_1, \dots, h_n)$. Then writing $K = X[k]X^*$ and $H = Y^*[h]Y$ with $h_1 k_1 \neq 0$, we have

$$\begin{aligned} W(CU^*AU) &= \lambda\mu W(X[k]X^*U^*Y^*[h]YU) \\ &= \lambda\mu W([k](YUX)^*[h](YUX)) \\ &= \lambda\mu W([k]R^*[h]R) \subseteq \mathbb{R}, \end{aligned}$$

and so it follows that $\lambda\mu k_1 h_1 \in \mathbb{R}$. Thus $\lambda\mu \in \mathbb{R}$ and so

$$W(KU^*HU) \subset \mathbb{R},$$

which implies that KU^*HU is hermitian for all U, i.e.,

$$KU^*HU = U^*HUK. \tag{80}$$

As we indicated above, both H and K may be assumed diagonal in (80). Since we are assuming that neither H nor K is scalar, we may also assume that

$$H = \begin{bmatrix} h_1 & 0 \\ 0 & h_2 \end{bmatrix} \dot{+} H_1,$$

$$K = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix} \dot{+} K_1,$$

in which $(h_1 - h_2)(k_1 - k_2) \neq 0$ and H_1 and K_1 are $(n-2)$-square diagonal matrices. If we take U to be

$$U = \begin{bmatrix} \bar{\xi} & \eta \\ -\bar{\eta} & \xi \end{bmatrix} \dot{+} I_{n-2}, \qquad |\xi|^2 + |\eta|^2 = 1,$$

and apply (80), we conclude that

$$\xi\eta(h_1 - h_2)(k_1 - k_2) = 0.$$

But $\xi\eta$ can be chosen not 0, and hence the assumption that $(h_1 - h_2)(k_1 - k_2) \neq 0$ (i.e., that both H and K are not scalar matrices) is contradicted. Thus at least one of H and K is scalar, and the proof is complete. ∎

REFERENCES

1 M. Goldberg and E. G. Straus, Elementary inclusion relations for generalized numerical ranges, *Linear Algebra Appl.* 18:1–24 (1977).

2 M. Goldberg and E. G. Straus, Norm properties of c-numerical radii, *Linear Algebra Appl.* 24:113–131 (1979).

3 M. Goldberg, On certain finite dimensional numerical ranges and numerical radii, *Linear and Multilinear Algebra* 7:329–342 (1979).

4 F. Hausdorff, Der Wertvorrat einer Bilinearform, *Math. Z.* 3:314–316 (1919).
5 M. Marcus and B. N. Moyls, Field convexity of a square matrix, *Proc. Amer. Math. Soc.* 6:981–985 (1955).
6 M. Marcus and A. J. Goldman, Convexity of the field of a linear transformation, *Canad. Math. Bull.* 2:15–18 (1959).
7 M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Prindle, Weber, & Schmidt, Boston, 1964.
8 M. Marcus, B. N. Moyls, and I. Filippenko, Normality and the higher numerical range, *Canad. J. Math.* 30:419–430 (1978).
9 M. Marcus and I. Filippenko, Linear operators preserving the decomposable numerical range, *Linear and Multilinear Algebra* 7:23–36 (1979).
10 M. Marcus and M. Sandy, Three elementary proofs of the Goldberg-Straus theorem on numerical radii, *Linear and Multilinear Algebra* 11:243–252 (1982).
11 F. D. Murnaghan, On the field of values of a square matrix, *Proc. Nat. Acad. Sci. U.S.A.* 18:246–248 (1932).
12 O. Toeplitz, Das Algebraischen Analogen zu einem Satze von Fejer, *Math. Z.* 2:187–197 (1918).
13 R. Westwick, A theorem on numerical range, *Linear and Multilinear Algebra* 2:311–315 (1975).

Received 12 December 1983; revised 22 February 1984