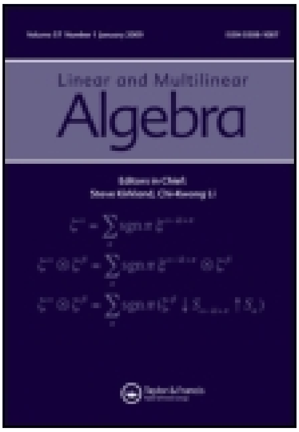

Linear and Multilinear Algebra
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/glma20

Numerical Range of a Normal Compression

Hwa-Long Gau [a] & Pei Yuan Wu [b]
[a] Department of Mathematics , National Central University , Chung-Li 320, Taiwan
[b] Department of Applied Mathematics , National Chiao Tung University , Hsinchu 300, Taiwan
Published online: 31 Aug 2006.

Linear and Multilinear Algebra, 2004, Vol. 52, Nos. 3–4, pp. 195–201

Numerical Range of a Normal Compression

HWA-LONG GAU[a,*] and PEI YUAN WU[b,†]

[a]*Department of Mathematics, National Central University, Chung-Li 320, Taiwan;*
[b]*Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan*

(Received 25 September 2002)

Let N be an n-by-n diagonal matrix whose distinct eigenvalues form corners of their convex hull, let x be a vector in $\mathbb{C}^n$ with nonzero components, and let A be the compression of N to the orthogonal complement of x. In this article, we study the properties of the eigenvalues and the numerical range of A and show that in many ways they are analogous to the ones for unitary N. The approach via the diagonal form of N yields a much simpler proof for some of the main results in this area.

Keywords: Numerical range; Normal compression

AMS Subject Classifications: 15A18; 15A60

The purpose of this article is to study the properties of the eigenvalues and numerical ranges of a class of matrices defined in the following way. Let

$$N = \operatorname{diag}(a_1, \ldots, a_n) \tag{1a}$$

be an n-by-n diagonal matrix with the a_js distinct such that each is a corner of the convex hull they generate. If

$$x = (x_1, \ldots, x_n)^T \tag{1b}$$

is a unit vector in $\mathbb{C}^n$ with $x_j \neq 0$ for all j, then let K be the orthogonal complement of the one-dimensional subspace generated by x. Finally, let A be the *compression* of N onto K, that is,

$$A = PN|_K, \tag{1c}$$

where P denotes the orthogonal projection from $\mathbb{C}^n$ onto K. In this case, we also say that N is a *dilation* of A. When N is unitary, such linear transformations A give exactly the $\mathcal{S}_{n-1}$-matrices or UB-matrices whose numerical ranges were studied in [3, 4, 6–9] in

*Corresponding author. E-mail: hlgau@math.ncu.edu.tw
†E-mail: pywu@math.nctu.edu.tw

ISSN 0308-1087 print: ISSN 0000-000X online/01/010000-00 © 2004 Taylor & Francis Ltd
DOI: 10.1080/0308108031000123659

recent years. Among other things, it was shown that their numerical ranges enjoy the Poncelet property (cf. [6, Theorem 1] or [3, Theorem 2.1]). In this article, we generalize results in this area from the unitary to the normal case. This is not merely a generalization for its own sake. By considering N in its diagonal form, we are able to adopt a simpler approach to prove many of the main theorems in this area in their generalized form. For this, we are indebted to Mirman who started his investigations using (more or less) the expression in (1c) and obtained many interesting results (cf. [6–9]). Our purpose here is to present a more streamlined and friendly line of arguments for the normal-compression case.

Our first theorem deals with the relation between the eigenvalues of A and N. For an n-by-n matrix X, let p_X denote its characteristic polynomial: $p_X(z) = \det(zI_n - X)$.

THEOREM 1 *If N is the diagonal matrix* $\text{diag}(a_1, \ldots, a_n)$ *and A is the compression of N to the orthogonal complement K of the unit vector $x = (x_1, \ldots, x_n)^T$ in $\mathbb{C}^n$, then the characteristic polynomial p_A of A is given by*

$$\sum_{j=1}^{n} |x_j|^2 (z - a_1) \ldots (\widehat{z - a_j}) \ldots (z - a_n), \tag{2}$$

where the hat "$\wedge$" over $z - a_j$ indicates that $z - a_j$ is absent from the product.

The assertion here can be deduced from a corresponding one on the rank-one perturbation.

LEMMA 2 *If $N = \text{diag}\,(a_1, \ldots, a_n)$ and $B = N + xy^*$, where $x = (x_1, \ldots, x_n)^T$ and $y = (y_1, \ldots, y_n)^T$, then the characteristic polynomials of N and B are related by*

$$p_B(z) = p_N(z) - \sum_{j=1}^{n} x_j \overline{y_j} (z - a_1) \ldots (\widehat{z - a_j}) \ldots (z - a_n). \tag{3}$$

This lemma is a special case of [2, Theorem 2]. We include a proof here for completeness.

Proof of Lemma 2 Let

$$F = xy^* = \begin{bmatrix} x_1\overline{y_1} & \cdots & x_1\overline{y_n} \\ \vdots & & \vdots \\ x_n\overline{y_1} & \cdots & x_n\overline{y_n} \end{bmatrix}.$$

Then

$$\begin{aligned} p_B(z) &= \det \begin{bmatrix} z - a_1 - x_1\overline{y_1} & -x_1\overline{y_2} & \cdots & -x_1\overline{y_n} \\ -x_2\overline{y_1} & z - a_2 - x_2\overline{y_2} & \cdots & -x_2\overline{y_n} \\ \vdots & \vdots & & \vdots \\ -x_n\overline{y_1} & -x_n\overline{y_2} & \cdots & z - a_n - x_n\overline{y_n} \end{bmatrix} \\ &= \sum_J (-1)^{\#J^c} \det(F|_{J^c}) \prod_{j \in J} (z - a_j), \end{aligned}$$

the summation being taken over all subsets J of $\{1, \ldots, n\}$ (including the empty set), where $\#J^c$ denotes the number of elements in the complement J^c of J in $\{1, \ldots, n\}$

and $\det(F|_{J^c})$ is the determinant of the submatrix $F|_{J^c}$ of F obtained by deleting the rows and columns indexed by integers in J. If $\#J \le n-2$, then, since $\operatorname{rank} F \le 1$ and the submatrix $F|_{J^c}$ has size at least two, we have $\operatorname{rank}(F|_{J^c}) \le 1$ and hence $\det(F|_{J^c}) = 0$. On the other hand, if $\#J = n-1$, say, $J = \{1, \ldots, \widehat{j}, \ldots, n\}$, then $\det(F|_{J^c}) = x_j\overline{y_j}$, and if $\#J = n$, then $\det(F|_{J^c}) = 1$. We conclude from the above expression for p_B that (3) holds. ■

Proof of Theorem 1 Letting $B = (I_n - xx^*)N$, we first show that $p_B(z) = zp_A(z)$. Indeed, since $I_n - xx^*$ is the orthogonal projection onto the $(n-1)$-dimensional subspace K of $\mathbb{C}^n$, there is a unitary matrix U such that $U^*(I_n - xx^*)U = \operatorname{diag}(1, \ldots, 1, 0)$. Then

$$\begin{aligned} U^*BU &= [U^*(I_n - xx^*)U]\,(U^*NU) \\ &= \operatorname{diag}(1,\ldots,1,0)\,(U^*NU) \\ &= \begin{bmatrix} A' & v \\ 0 & 0 \end{bmatrix}, \end{aligned}$$

where A' and v are, respectively, $(n-1)$-by-$(n-1)$ and $(n-1)$-by-1 matrices. Hence

$$p_B(z) = P_{U^*BU}(z) = zp_{A'}(z).$$

On the other hand, if $C = (I_n - xx^*)N(I_n - xx^*)$, then

$$\begin{aligned} U^*CU &= [U^*(I_n - xx^*)U](U^*NU)[U^*(I_n - xx^*)U] \\ &= \operatorname{diag}(1,\ldots,1,0)(U^*NU)\operatorname{diag}(1,\ldots,1,0) \\ &= \begin{bmatrix} A' & 0 \\ 0 & 0 \end{bmatrix}, \end{aligned}$$

which shows that A' is a matrix representation of A. Thus $p_B(z) = zp_A(z)$ as required. Since

$$B = N - xx^*N = N + xy^*,$$

where $y = (-x_1\overline{a_1}, \ldots, -x_n\overline{a_n})^T$, we may apply Lemma 2 to obtain

$$\begin{aligned} p_B(z) &= p_N(z) + \sum_{j=1}^{n} x_j(\overline{x_j}a_j)(z-a_1)\ldots\widehat{(z-a_j)}\ldots(z-a_n) \\ &= \left(\sum_j |x_j|^2\right)(z-a_1)\ldots(z-a_n) + \sum_j |x_j|^2 a_j(z-a_1)\ldots\widehat{(z-a_j)}\ldots(z-a_n) \\ &= z\sum_j |x_j|^2(z-a_1)\ldots\widehat{(z-a_j)}\ldots(z-a_n). \end{aligned}$$

(2) then follows immediately. ■

Several remarks are in order. Firstly, (2) should be contrasted, when N is unitary, with the condition

$$p_N(z) = z\prod_{j=1}^{n-1}(z-b_j) - e^{i\theta}\prod_{j=1}^{n-1}(1-\overline{b_j}z) \tag{4}$$

expressing the characteristic polynomial p_N of N in terms of the eigenvalues $b_1, \ldots, b_{n-1}$ of A and the parameter $e^{i\theta}$ (cf. [4, Lemma 2.4]). (2) is obtained from the diagonal form of N and depends on the unit vector x while (4) is from the upper-triangular form of A and depends on the parameter $e^{i\theta}$. A relation involving only eigenvalues of N and A can be easily derived from (4) (cf. [4, Theorem 2.5]). But one from (2) seems more complicated. They are presumably equivalent to each other.

Secondly, if, besides the conditions of Theorem 1, we further assume that the a_js are distinct and the x_js are all nonzero, then the eigenvalues of A are exactly those complex numbers z satisfying

$$\sum_{j=1}^{n} \frac{|x_j|^2}{z - a_j} = 0.$$

This is an easy consequence of (2) and is an alternative form in the literature for eigenvalues of A in this situation (cf. [6, Theorem 6 (7)]).

Finally, if the a_js are the zeros of a degree-n polynomial p and $x = (1/\sqrt{n}, \ldots, 1/\sqrt{n})^T$, then the eigenvalues of the corresponding A are exactly the zeros of the derivative p' of p. Indeed, this is because

$$\begin{aligned} p_A(z) &= \frac{1}{n}\sum_{j=1}^{n}(z - a_1)\ldots(\widehat{z - a_j})\ldots(z - a_n) \\ &= \frac{1}{n} p'(z) \end{aligned}$$

by (2).

We now move from eigenvalues to the numerical range. Recall that the *numerical range* $W(B)$ of any n-by-n matrix B is the subset $\{\langle Bx, x\rangle : x \in \mathbb{C}^n, \|x\| = 1\}$ of the plane, where $\langle\cdot,\cdot\rangle$ denotes the usual inner product in $\mathbb{C}^n$. Properties of the numerical range can be found in [5, Chap. 1]. The next theorem gives properties of the numerical range of the linear transformation A defined in (1c). Its first part is essentially in [9] and [1, Theorem 1].

THEOREM 3. *Let N, x and A be as in* (1a), (1b) *and* (1c), *respectively, and assume that the a_js are consecutive corners of their convex hull. Then $W(A) \cap (a_j, a_{j+1})$ is a singleton, say, $\{c_j\}$ for each $j = 1, \ldots, n$ $(a_{n+1} \equiv a_1)$ and they satisfy*

$$\prod_{j=1}^{n} |c_j - a_j| = \prod_{j=1}^{n} |c_j - a_{j+1}|.$$

Conversely, if N is as given in (1a) *with a_js the consecutive corners of their convex hull and if $t_1, \ldots, t_n$ in $(0, \infty)$ are such that $t_1 \ldots t_n = 1$, then there is a unit vector $x = (x_1, \ldots, x_n)^T$ with $x_j \neq 0$ for all j such that the corresponding A in* (1c) *has its numerical range $W(A)$ intersecting (a_j, a_{j+1}) at a single point c_j which satisfies $|c_j - a_j|/|c_j - a_{j+1}| = t_j$ for each j.*

Proof For $j = 1, \ldots, n$, let e_j be the vector in $\mathbb{C}^n$ whose jth component is one and other components are all zero, and let $e_{n+1} = e_1$. If

$$y_j = \frac{1}{\left(|x_j|^2 + |x_{j+1}|^2\right)^{1/2}}(\overline{x_{j+1}}e_j - \overline{x_j}e_{j+1}), \quad j = 1, \ldots, n,$$

then y_j is a unit vector orthogonal to x. Hence

$$\begin{aligned} c_j &\equiv \langle Ay_j, y_j\rangle = \langle Ny_j, y_j\rangle \\ &= \frac{1}{|x_j|^2 + |x_{j+1}|^2}(a_j|x_{j+1}|^2 + a_{j+1}|x_j|^2), \end{aligned}$$

which shows that c_j is in $W(A) \cap (a_j, a_{j+1})$. To prove that c_j is the only point in this intersection, let d belong to $W(A) \cap (a_j, a_{j+1})$. Then $d = \langle Au, u\rangle$ for some unit vector $u = (u_1, \ldots, u_n)^T$ in K, the orthogonal complement of x. We have

$$d = \langle Nu, u\rangle = \sum_{k=1}^{n} a_k|u_k|^2$$

in (a_j, a_{j+1}). Since the a_ks are corners of their convex hull, the convex combination of d in terms of the a_ks has unique coefficients. Hence we must have $u_k = 0$ for all $k \neq j, j+1$. Thus u and y_j are both in the two-dimensional subspace $M_j \equiv \vee\{e_j, e_{j+1}\}$. If u and y_j are linearly independent, then they will span M_j. In this case, e_j will be a linear combination of u and y_j and hence is orthogonal to x, which leads to $x_j = 0$, a contradiction. Therefore, u and y_j must be linearly dependent. This implies that

$$d = \langle Au, u\rangle = \langle Ay_j, y_j\rangle = c_j.$$

Hence $W(A) \cap (a_j, a_{j+1}) = \{c_j\}$ and

$$\prod_{j=1}^{n} |c_j - a_j| = \prod_{j=1}^{n} \frac{|a_{j+1} - a_j||x_j|^2}{|x_j|^2 + |x_{j+1}|^2} = \prod_{j=1}^{n} |c_j - a_{j+1}|.$$

To prove the converse, let

$$x_j = \frac{(t_j \ldots t_n)^{1/2}}{\left(\sum_{k=1}^{n} t_k \ldots t_n\right)^{1/2}}, \quad j = 1, \ldots, n.$$

Then $x_j \neq 0$ for all j and $\sum_j |x_j|^2 = 1$. Let $x = (x_1, \ldots, x_n)^T$ and let A be the corresponding compression of N as in (1c). Then, as in the proof of the first part, the numerical range $W(A)$ intersects (a_j, a_{j+1}) at the point

$$c_j = \frac{1}{|x_j|^2 + |x_{j+1}|^2}(a_j|x_{j+1}|^2 + a_{j+1}|x_j|^2)$$

and thus

$$\frac{|c_j - a_j|}{|c_j - a_{j+1}|} = \frac{|x_j|^2}{|x_{j+1}|^2} = \frac{t_j \dots t_n}{t_{j+1} \dots t_n} = t_j$$

for all j. This completes the proof. ■

Again, in the above situation, if the a_js are the zeros of a degree-n polynomial p and $x = (1/\sqrt{n}, \dots, 1/\sqrt{n})^T$, then the intersection point of $W(A)$ and (a_j, a_{j+1}) is easily seen to be $(a_j + a_{j+1})/2$. This much easier proof of [4, Theorem 2.1] is due to Mirman.

We conclude this article with a uniqueness result that the compression A in (1c) is determined, up to unitary equivalence, by the intersection points of $W(A)$ and (a_j, a_{j+1}). This is a generalization of [3, Theorem 3.2].

THEOREM 4 *Let N be as in* (1a) *with a_js the consecutive corners of their convex hull, let $x = (x_1, \dots, x_n)^T$ and $y = (y_1, \dots, y_n)^T$ be unit vectors in $\mathbb{C}^n$ with nonzero x_js and y_js, and let A and B be the compressions of N to the orthogonal complements of x and y, respectively. Then the following are equivalent*:

(a) *A and B are unitarily equivalent*;
(b) $W(A) = W(B)$;
(c) $W(A) \cap (a_j, a_{j+1}) = W(B) \cap (a_j, a_{j+1})$ *for all j*.

Proof We need only prove that (c) ⇒ (a). By Theorem 3, the common intersection $W(A) \cap (a_j, a_{j+1}) = W(B) \cap (a_j, a_{j+1})$ is a singleton, say, $\{c_j\}$. Let $t_j = |c_j - a_j|/|c_j - a_{j+1}|$ for each j. As proved in Theorem 3, we have

$$t_j = \frac{|x_j|^2}{|x_{j+1}|^2} = \frac{|y_j|^2}{|y_{j+1}|^2}.$$

It follows from the first equality that $|x_j|^2 = t_j |x_{j+1}|^2$ and thus

$$|x_j|^2 = t_j \dots t_n |x_{n+1}|^2 = t_j \dots t_n |x_1|^2$$

by induction. Since x is a unit vector, we obtain

$$1 = \sum_{j=1}^{n} |x_j|^2 = \left(\sum_{j=1}^{n} t_j \dots t_n \right) |x_1|^2$$

and therefore

$$|x_1| = \frac{1}{(\sum_{j=1}^{n} t_j \dots t_n)^{1/2}}.$$

The same goes for the y_js and hence $|x_j| = |y_j|$ for all j. Thus $x_j = y_j e^{i\theta_j}$ for some real θ_j. If $U = \mathrm{diag}\,(e^{i\theta_1}, \dots, e^{i\theta_n})$, then U is a unitary matrix commuting with N and $Uy = x$. Consider the orthogonal decompositions $\mathbb{C}^n = K \oplus K^\perp$ and $\mathbb{C}^n = L \oplus L^\perp$, where

K and L are the orthogonal complements of x and y, respectively. Then N is unitarily equivalent to matrices of the form

$$\begin{bmatrix} A & * \\ * & * \end{bmatrix} \text{ on } K \oplus K^{\perp} \text{ and } \begin{bmatrix} B & * \\ * & * \end{bmatrix} \text{ on } L \oplus L^{\perp}.$$

Moreover, U is unitarily equivalent to a matrix of the form

$$\begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \text{ from } L \oplus L^{\perp} \text{ onto } K \oplus K^{\perp}.$$

The commutativity of U and N implies that

$$\begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \begin{bmatrix} B & * \\ * & * \end{bmatrix} = \begin{bmatrix} A & * \\ * & * \end{bmatrix} \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix}$$

and thus $U_1 B = A U_1$, which shows that A is unitarily equivalent to B. ■

Now the final remark. As we have seen before, Theorems 3 and 4 generalize the corresponding results for $\mathcal{S}_{n-1}$-matrices. However, for $n = 2, 3$ or 4, they can also be obtained from the latter. This is because by a reduction via an affine transformation and some computations with the quadratic polynomial in two variables we can show that *four distinct points on the plane lie on an ellipse if and only if each is a corner of their convex hull*. When n is at most four, we then apply an affine transformation which maps the unit circle to an ellipse passing the eigenvalues of N to the corresponding results for $\mathcal{S}_{n-1}$-matrices [3, Theorems 3.1 and 3.2] to obtain Theorems 3 and 4 (since the interscribing property is preserved under this process). We leave out the details.

Acknowledgment

This research was supported by the National Science Council of the Republic of China under research projects NSC-91-2115-M-008-011 and NSC-91-2115-M-009-007 of the respective authors.

References

[1] M. Adam and J. Maroulas (2002). On compressions of normal matrices. *Linear Algebra Appl.*, **341**, 403–418.

[2] J. Anderson (1996). A secular equation for the eigenvalues of a diagonal matrix perturbation. *Linear Algebra Appl.*, **246**, 49–70.

[3] H.-L. Gau and P.Y. Wu (1998). Numerical range of $S(\phi)$. *Linear and Multilinear Algebra*, **45**, 49–73.

[4] H.-L. Gau and P.Y. Wu (1999). Lucas' theorem refined. *Linear and Multilinear Algebra*, **45**, 359–373.

[5] R.A. Horn and C.R. Johnson (1991). *Topics in Matrix Analysis*. Cambridge Univ. Press, Cambridge.

[6] B. Mirman (1998). Numerical ranges and Poncelet curves. *Linear Algebra Appl.*, **281**, 59–85.

[7] B. Mirman (2003). UB-matrices and conditions for Poncelet polygon to be closed. *Linear Algebra Appl.*, **360**, 123–150.

[8] B. Mirman, V. Borovikov, L. Ladyzhensky and R. Vinograd (2001). Numerical ranges, Poncelet curves, invariant measures. *Linear Algebra Appl.*, **329**, 61–75.

[9] B. Mirman and P.Y. Wu. Matrix interpretation of Marden's proof of Siebeck's theorem (preprint).