SIAM J. MATRIX ANAL. APPL.
Vol. 21, No. 1, pp. 230–244


# COMPRESSIONS AND DILATIONS OF NUMERICAL RANGES*

J. MAROULAS† AND M. ADAM†

**Abstract.** Inner and outer approximation of numerical ranges of $n \times n$ complex matrices and matrix polynomials is investigated in this paper, which is based on the numerical ranges of matrices of smaller or double dimensions.

**Key words.** numerical range, matrix polynomials

**AMS subject classifications.** 15A60

**PII.** S0895479898340524

## 1. Introduction.

Let $M_n$ be the algebra of all $n \times n$ complex matrices. For a matrix $A \in M_n$, the *numerical range* $NR[A]$, also known as the *field of values*, is the set of complex numbers

$$NR[A] = \{\langle Ax, x\rangle : \quad x \in C^n, \quad \|x\| = 1\}. \tag{1.1}$$

The radius

$$r(A) = \max\{|z| : \quad z \in NR[A]\}$$

of the smallest circle $|z| = r(A)$ that encloses $NR[A]$ is called the *numerical radius.* The usefulness of the numerical range and numerical radii is well known. As an extensive background for this active research topic we refer to [HJ] and [AL]. Replacing in (1.1) the Euclidean inner product with the *indefinite scalar product* on $\mathbf{C}^n$, it is known [GLR] that there exists an invertible indefinite hermitian matrix $S$ such that $\langle x, y\rangle_S = \langle Sx, y\rangle$. Hence, we obtain the *S-numerical range* of $A$:

$$W_S(A) = \left\{ \frac{\langle Ax, x\rangle_S}{\langle x, x\rangle_S} : \quad x \in \mathbf{C}^n, \quad \langle x, x\rangle_S \neq 0 \right\} = W_S^+(A) \bigcup W_{-S}^+(A), \tag{1.2}$$

where

$$W_S^+ = \{\langle Ax, x\rangle_S : \quad x \in \mathbf{C}^n, \quad \langle x, x\rangle_S = 1\}. \tag{1.3}$$

In particular, the set $W_S^+(A)$ is called the *positive S-numerical range.* The $S$-numerical ranges generalize the classical numerical range, and some properties of the $NR[A]$ can be extended to $W_S(A)$. In [B] it is proved that $W_S^+(A)$ is a convex set and also that the closure of set $W_S(A)$ contains all eigenvalues of $A$ if $A$ is positive definite. Moreover, we can readily verify the following:

(i) $W_S^+(A + kI) = W_S^+(A) + k, \quad k \in \mathbf{C}$.
(ii) $W_S^+(kA) = kW_S^+(A), \quad k \in \mathbf{C}$.
(iii) $W_S^+(A + B) \subseteq W_S^+(A) + W_S^+(B)$.
(iv) $W_S^+(A_1) \subseteq W_S^+(A), A_1$, a submatrix of $A$.

* Received by the editors June 4, 1998; accepted for publication (in revised form) by R. Bhatia November 6, 1998; published electronically September 22, 1999.
http://www.siam.org/journals/simax/21-1/34052.html

†Department of Mathematics, National Technical University, Zografou Campus, Athens 15780, Greece (maroulas@math.ntua.gr, maria@math.ntua.gr).

(v) $z \in W_S^+(A)$, $S$ real symmetric $\Rightarrow \bar{z} \in W_S^+(\bar{A})$.

A further study of both $W_S(A)$ and $W_S^+(A)$ has been presented by [LTU], where it is shown that $W_S(A)$ is always $p$-convex, i.e., for any distinct pair of points $z_1$, $z_2 \in W_S(A)$ either the closed line segment $\overline{z_1 z_2} \subset W_S(A)$ or the line $\{az_1 + (1-a)z_2 : a \le 0 \text{ or } 1 \le a\} \subset W_S(A)$. Clearly, if $S$ is positive definite, then

$$W_S(A) = W_S^+(A) = NR[S^{1/2}AS^{-1/2}].$$

The idea of the numerical range has also been extended to matrix polynomials [LR], [M]:

$$L(\lambda) = I\lambda^q + A_1\lambda^{q-1} + \cdots + A_q; \;\; A_i \in M_n,$$

where

$$NR[L(\lambda)] = \{\lambda : \; \langle L(\lambda)x, x\rangle = 0 \text{ for some } x \in \mathbf{C}^n\}. \tag{1.4}$$

$NR[L(\lambda)]$ has peculiar geometric properties with respect to the boundedness and connectedness of this set, which are important to the factorization of matrix polynomials [M], [MP1, MP2, MP3], [MMP], and [LMZ].

In this paper we continue our effort for further development of the subject and the study of related problems. In section 2 we express $NR[A]$ as the union of the numerical ranges of matrices of dimensions $k \times k$ for $2 \le k < n$. In this way, each set in the union can be considered as an *inner approximation* or *dilation* of $NR[A]$. This result generalizes the approach of Markus and Pesce [MP], where $NR[A]$ is the union of the numerical ranges of $2 \times 2$ matrices, i.e., of ellipsoid regions using only real orthogonal vectors. Taking advantage of the fact that $NR[A]$ and $NR[e^{2i\theta}\bar{A}]$ are symmetric with respect to the straight line $y = (\tan\theta)x$, we have set the convex hull of $NR[A] \bigcup NR[e^{2i\theta}\bar{A}]$ equal to the numerical range of a suitable matrix. This result for $\theta = 0$ leads to the convenient equality

$$\text{Conv.hull}\left(NR[A] \cup NR[\bar{A}]\right) = NR\left(\begin{bmatrix} M & N \\ -N & M \end{bmatrix}\right),$$

where $M, N$ are real matrices defined by $A = M + iN$. Therefore, $NR[A]$ is presented as the intersection of numerical ranges as the line $y = (\tan\theta)x$ rotates around the origin for $0 \le \theta \le \pi$. The last equality is generalized further for the matrix

$$C = \begin{bmatrix} A & B \\ -\bar{B} & \bar{A} \end{bmatrix},$$

which, as is known, corresponds to a quartenionic matrix. Clearly, the outer set Conv.hull($NR[A] \cup NR[\bar{A}]$) can be considered as a *compression* of $NR[A]$. Some additional comments on projections on the axes, on the joint numerical range [HJ], and on the numerical radius have been presented. The MATLAB procedure is exposed and examples are given.

In the third section we refer to $W_S^+(A)$. We show that for any indefinite hermitian matrix $S$,

$$NR[A] \cap W_S^+(A) \ne \emptyset,$$

and also that

$$W^{+}_{I_2\otimes S}(A\oplus B) = \text{Conv.hull}(W^{+}_{S}(A)\cup W^{+}_{S}(B)).$$

This equality clarifies Conv.hull($W^{+}_{S}(A)\cup W^{+}_{S}(e^{2i\theta}\bar{A})$) as the $(I_2\otimes S)$-numerical range of a matrix depending on $A$.

Finally, in the fourth section, we present two properties of $NR[L(\lambda)]$. In particular, the first one refers to the reduction of this set to $NR[M(\lambda)]$, where the $k\times k$ $(k\geq 2)$ matrix polynomial $M(\lambda)$ depends on $L(\lambda)$.

**2. The numerical range of a matrix.** An approximation or a *dilation* of $NR[A]$ in (1.1) is presented as a first statement.

THEOREM 2.1. *For a prefix number $k<n$,*

$$NR[A] = \bigcup_{\xi_1,\dots,\xi_k} NR\left(\begin{bmatrix} \xi_1^*A\xi_1 & \dots & \xi_1^*A\xi_k \\ \vdots & & \vdots \\ \xi_k^*A\xi_1 & \dots & \xi_k^*A\xi_k \end{bmatrix}\right), \tag{2.1}$$

*where $\xi_1,\dots,\xi_k$ run over all sets by $k$ orthonormal vectors of $\mathbf{C}^n$.*

*Proof.* Any vector $x\in\mathbf{C}^n$ belongs to a $k$-dimensional subspace $\mathbf{E}\subset\mathbf{C}^n$. Let $\xi_1,\dots,\xi_k$ be an orthonormal basis of $\mathbf{E}$ such that $\xi_i\in\mathbf{C}^n$. Then

$$x = [\xi_1\dots\xi_k]\,\omega,$$

where $\omega\in\mathbf{C}^k$. Clearly, $\omega$ is a unit vector if and only if $x$ is a unit, and due to

$$x^*Ax = \omega^*\begin{bmatrix}\xi_1^*\\ \vdots\\ \xi_k^*\end{bmatrix} A\,[\xi_1\dots\xi_k]\,\omega = \omega^*\begin{bmatrix} \xi_1^*A\xi_1 & \dots & \xi_1^*A\xi_k \\ \vdots & & \vdots \\ \xi_k^*A\xi_1 & \dots & \xi_k^*A\xi_k \end{bmatrix}\omega,$$

we verify the relationship (2.1). □

For $k=2$, the interior of $NR[A]$ is covered by the ellipses

$$NR\left(\begin{bmatrix} \xi_1^*A\xi_1 & \xi_1^*A\xi_2 \\ \xi_2^*A\xi_1 & \xi_2^*A\xi_2 \end{bmatrix}\right),$$

where $\{\xi_1,\ \xi_2\}$ is any pair of orthonormal vectors. The estimation of $NR[A]$ clearly is realized easily if $k>2$, since the fields of values in the left part of (2.1) are convex sets and even they occupy a more extensive area of $NR[A]$ as $k$ increases. This is illustrated in example 1. In particular, for $k=2$ and for $\xi_1,\ \xi_2$ real orthogonal vectors, (2.1) has been presented in [MP] using a different approach.

A similar approximation for the $q$-numerical range of $A$:

$$NR_q[A] = \{\,y^*Ax \ \ :\ \|x\|=\|y\|=1,\ y^*x=q\},\quad q\in[0,1],$$

is given by the formula:

$$NR_q[A] = \bigcup_{\xi_1,\xi_2} NR_q\left(\begin{bmatrix} \xi_1^*A\xi_1 & \xi_1^*A\xi_2 \\ \xi_2^*A\xi_1 & \xi_2^*A\xi_2 \end{bmatrix}\right),$$

since for $x \in \mathbf{\Delta} = \text{span}\{\xi_1, \xi_2\}$, there exists $y \in \mathbf{\Delta}$ such that $y^*x = q$.

For the *compression* of $NR[A]$ we state the following theorem.

THEOREM 2.2. *For any matrix* $A$,

$$\text{(2.2) Conv.hull}(NR[A] \cup NR[e^{2i\theta}\bar{A}]) = NR\left(\frac{1}{2}\begin{bmatrix} A + e^{2i\theta}\bar{A} & -i(A - e^{2i\theta}\bar{A}) \\ i(A - e^{2i\theta}\bar{A}) & A + e^{2i\theta}\bar{A} \end{bmatrix}\right),$$

*where* $0 \le \theta \le \pi$.

*Proof.* Since

$$\text{Conv.hull}(NR[A] \cup NR[e^{2i\theta}\bar{A}]) = NR\left(\begin{bmatrix} A & O \\ O & e^{2i\theta}\bar{A} \end{bmatrix}\right),$$

we have to show that

$$NR\left(\frac{1}{2}\begin{bmatrix} A + e^{2i\theta}\bar{A} & -i(A - e^{2i\theta}\bar{A}) \\ i(A - e^{2i\theta}\bar{A}) & A + e^{2i\theta}\bar{A} \end{bmatrix}\right) = NR\left(\begin{bmatrix} A & O \\ O & e^{2i\theta}\bar{A} \end{bmatrix}\right).$$

In fact, for a unit vector $x \in \mathbf{C}^{2n}$ and for $A = M + iN$, where $M, N \in \mathbf{R}_{n\times n}$, we have

$$x^*\begin{bmatrix} A & O \\ O & e^{2i\theta}\bar{A} \end{bmatrix}x = \frac{1}{2}x^*\begin{bmatrix} I & O \\ O & e^{2i\theta}I \end{bmatrix}\begin{bmatrix} I & -iI \\ -iI & I \end{bmatrix}\begin{bmatrix} M & N \\ -N & M \end{bmatrix}\begin{bmatrix} I & iI \\ iI & I \end{bmatrix}x$$

$$= \frac{1}{4}x^*\begin{bmatrix} I & -iI \\ -iI & I \end{bmatrix}\begin{bmatrix} (1+e^{2i\theta})I & -i(1-e^{2i\theta})I \\ i(1-e^{2i\theta})I & (1+e^{2i\theta})I \end{bmatrix}$$

$$\times \begin{bmatrix} M & N \\ -N & M \end{bmatrix}\begin{bmatrix} I & iI \\ iI & I \end{bmatrix}x$$

$$= \frac{1}{2}x^*\left(\begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} \otimes I\right)\begin{bmatrix} \frac{A+e^{2i\theta}\bar{A}}{2} & \frac{-i(A-e^{2i\theta}\bar{A})}{2} \\ \frac{i(A-e^{2i\theta}\bar{A})}{2} & \frac{A+e^{2i\theta}\bar{A}}{2} \end{bmatrix}$$

$$\times \left(\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \otimes I\right)x$$

$$= y^*\begin{bmatrix} \frac{A+e^{2i\theta}\bar{A}}{2} & \frac{-i(A-e^{2i\theta}\bar{A})}{2} \\ \frac{i(A-e^{2i\theta}\bar{A})}{2} & \frac{A+e^{2i\theta}\bar{A}}{2} \end{bmatrix}y,$$

where $y = \frac{1}{\sqrt{2}}x^*\left(\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \otimes I\right)x$. Since

$$y^*y = \frac{1}{2}x^*\left(\begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} \otimes I\right)\left(\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \otimes I\right)x = x^*x = 1,$$

the relationship (2.2) is verified readily. □

COROLLARY 2.3. *For any matrix* $A = M + iN$, *where* $M, N \in \mathbf{R}_{n\times n}$,

$$\text{(2.3)} \qquad \text{Conv.hull}\ (NR[A] \cup NR[\bar{A}]) = NR\left(\begin{bmatrix} M & N \\ -N & M \end{bmatrix}\right).$$

*Proof.* Equation (2.3) comes from (2.2) with $\theta = 0$. □

From Corollary 2.3, we have the following comments.

(I) For any matrix $A = M + iN$, the numerical range of a $2n \times 2n$ matrix $\begin{bmatrix} M & N \\ -N & M \end{bmatrix}$ is symmetric to the real axis.

(II)

$$NR\left(\begin{bmatrix} M & -N \\ N & M \end{bmatrix}\right) = NR\left(\begin{bmatrix} M & N \\ -N & M \end{bmatrix}\right). \tag{2.4}$$

This is evident, due to

$$\begin{bmatrix} M & -N \\ N & M \end{bmatrix} = \begin{bmatrix} O & I \\ I & O \end{bmatrix} \begin{bmatrix} M & N \\ -N & M \end{bmatrix} \begin{bmatrix} O & I \\ I & O \end{bmatrix}.$$

(III)

$$\text{proj}_{ox} NR[A] = \text{proj}_{ox} NR\left(\begin{bmatrix} M & N \\ -N & M \end{bmatrix}\right). \tag{2.5}$$

In fact, by the symmetry of $NR[A]$ and $NR[\bar{A}]$ to the real axis, it is implied that the field Conv.hull($NR[A] \cup NR[\bar{A}]$) is also symmetric to the same axis. Hence, by (2.3) we obtain (2.5).

(IV)

$$\text{proj}_{oy} NR[A] = \text{proj}_{ox} NR\left(\begin{bmatrix} N & M \\ -M & N \end{bmatrix}\right). \tag{2.6}$$

By the equalities (2.4) and $\text{Im}(x^*Ax) = \text{Re}(x^*(-iA)x)$, (2.6) is obvious.

(V) The real field of values [HJ, p. 85]

$$JNR\left(\begin{bmatrix} M & -N \\ N & M \end{bmatrix}, \begin{bmatrix} N & M \\ -M & N \end{bmatrix}\right) \equiv NR[A]. \tag{2.7}$$

In fact, if we denote the unit vector $x = u + iv$, then by the relationship

$$x^*Ax = w^T \begin{bmatrix} M & -N \\ N & M \end{bmatrix} w + i\left(w^T \begin{bmatrix} N & M \\ -M & N \end{bmatrix} w\right),$$

where $w = \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbf{R}^{2n}$, we confirm the truth of the contention.

THEOREM 2.4. *For any matrix* $A \in \mathbf{C}_{n\times n}$,

$$NR[A] = \bigcap_{0\le\theta\le\pi} NR\left(\frac{1}{2}\begin{bmatrix} A + e^{2i\theta}\bar{A} & -i(A - e^{2i\theta}\bar{A}) \\ i(A - e^{2i\theta}\bar{A}) & A + e^{2i\theta}\bar{A} \end{bmatrix}\right). \tag{2.8}$$

*Proof.* Let $z \in NR[A]$. Then the complex number

$$w = e^{2i\theta}\bar{z} = \frac{1 + i\lambda}{1 - i\lambda}\bar{z},$$

where $\lambda = \tan\theta$, is symmetric in $z$ with respect to the straight line $y = \lambda x$. Since $w \in NR[e^{2i\theta}\bar{A}]$, the fields $NR[A]$ and $NR[e^{2i\theta}\bar{A}]$ are symmetric with respect to the line $y = \lambda x$, and hence by (2.2),

$$\bigcap_{0\le\theta\le\pi} NR\left(\frac{1}{2}\begin{bmatrix} A+e^{2i\theta}\bar{A} & -i(A-e^{2i\theta}\bar{A}) \\ i(A-e^{2i\theta}\bar{A}) & A+e^{2i\theta}\bar{A}\end{bmatrix}\right)$$

$$= \bigcap_{0\le\theta\le\pi} \text{Conv.hull}\left(NR[A] \cup NR\left[\frac{1+i\lambda}{1-i\lambda}\bar{A}\right]\right) = NR[A]. \qquad \square$$

Setting $e^{-i\theta}A = S_\theta + iT_\theta$, where $S_\theta, T_\theta$ are real matrices, by (2.8) we have directly the following corollary.

COROLLARY 2.5.

$$NR[A] = \bigcap_{0\le\theta\le\pi} NR\left(e^{i\theta}\begin{bmatrix} S_\theta & T_\theta \\ -T_\theta & S_\theta\end{bmatrix}\right). \tag{2.9}$$

Hence, by the theory of the numerical range of matrices [HJ], we can say that Theorems 2.1 and 2.4 give a better approximation of $NR[A]$ than the polygonal approximation. This happens because the boundary of the numerical ranges of matrices in (2.1) and also of those in (2.8) contain parts of $\partial NR[A]$ and not points, i.e., the vertices of polygons. Hence, the verification of $NR[A]$ by the intersection in (2.8), due to the convexity of $NR[A]$, needs a small number for $\theta$.

A further generalization of (2.3) is the investigation of a relationship between the numerical ranges of matrices:

$$C = \begin{bmatrix} A & B \\ -\bar{B} & \bar{A}\end{bmatrix} K = \begin{bmatrix} M & N & \vdots & U & V \\ -N & M & \vdots & -V & U \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -U & V & \vdots & M & -N \\ -V & -U & \vdots & N & M \end{bmatrix}, \tag{2.10}$$

where $A = M + iN$ and $B = U + iV$.

THEOREM 2.6. *For the matrices $C$ and $K$ in* (2.10) *it is true that* $NR[C] = NR[K]$.

*Proof.* Denote by $R$ the unitary matrix

$$R = \frac{1}{2}\begin{bmatrix} 1 & i & 1 & -i \\ i & 1 & -i & 1 \\ -i & 1 & i & 1 \\ -1 & i & -1 & -i \end{bmatrix} \otimes I.$$

Then

$$R^*KR = \text{diag}(C, \bar{C}),$$

and thus we obtain

$$NR[K] = NR[\text{diag}(C,\bar{C})] = \text{Conv.hull}\left(NR[C] \cup NR[\bar{C}]\right).$$

Moreover, by the equation

$$\begin{bmatrix} O & -I \\ I & O \end{bmatrix} C \begin{bmatrix} O & I \\ -I & O \end{bmatrix} = \bar{C},$$

we have $NR[C] = NR[\bar{C}]$ and consequently the equality of $NR[C]$ and $NR[K]$ is evident. □

In the special case where

$$C = \begin{bmatrix} p+iq & u+iv \\ -u+iv & p-iq \end{bmatrix}, \qquad p,q,u,v \in \mathbf{R},$$

we can check that $NR[C]$ is the interval $[p-i\sqrt{q^2+u^2+v^2},\ p+i\sqrt{q^2+u^2+v^2}]$.

For the *numerical radius* $r(A)$, defined by $r(A) = \max\{|z| : z \in NR[A]\}$, we verify the following theorem.

THEOREM 2.7.

$$r\left(\begin{bmatrix} \xi_1^* A\xi_1 & \dots & \xi_1^* A\xi_k \\ \vdots & & \vdots \\ \xi_k^* A\xi_1 & \dots & \xi_k^* A\xi_k \end{bmatrix}\right) \le r(A) = r\left(\begin{bmatrix} A & O \\ O & e^{2i\theta}\bar{A} \end{bmatrix}\right). \tag{2.11}$$

*Proof.* By the relationship (2.1), clearly the $NR$ of the matrix on the right is a subset of $NR[A]$, which implies the inequality in (2.11). For the equality, note that

$$r\left(\begin{bmatrix} A & O \\ O & e^{2i\theta}\bar{A} \end{bmatrix}\right) = \max\left(r(A),\ r(e^{2i\theta}\bar{A})\right) = r(A),$$

since $r(e^{2i\theta}\bar{A}) = |e^{2i\theta}|r(\bar{A}) = r(A)$. □

**MATLAB procedure.** The inner approximation of $NR[A]$, as it is expressed in Theorem 2.1, can be illustrated as follows:

Step 1. Introduce the matrix.

Step 2. For a prefix number $k < n$ introduce $k$ arbitrary linearly independent vectors of $\mathbf{C}^n$.

Step 3. Orthonormalize this set to $\xi_1, \dots, \xi_k$.

Step 4. Calculate the matrix $B = [\xi_i^* A\xi_j]$, $i,j = 1,2,\dots,k$.

Step 5. Determine the numerical range of $B$.

Step 6. Illustrate $NR[B]$.

Step 7. Repeat this procedure for some other set of vectors.

*Example* 1. Let

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

In Figure 2.1, the $NR[A]$ are defined as the union of the numerical ranges of $2\times 2$ and $3\times 3$ matrices, an application of (2.1) for $k=2$ and $k=3$.

As an implication of Theorem 2.4 we give the next example.

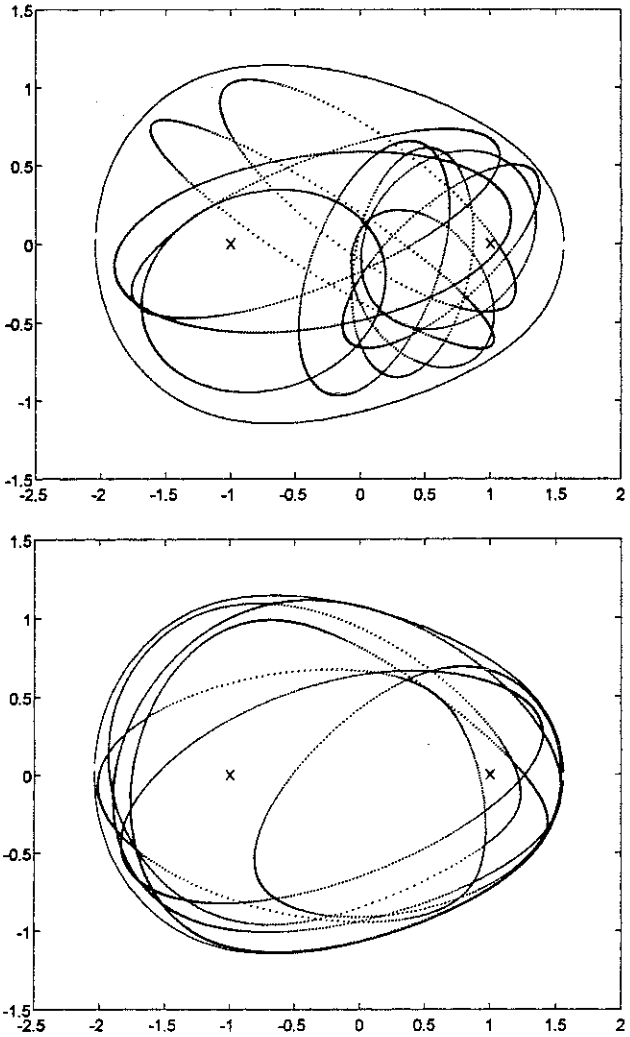


FIG. 2.1

*Example* 2. Let

$$A = \begin{bmatrix} 2-i & 1 \\ 0 & 1-i \end{bmatrix}.$$

Then for $\theta = 0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}$ we illustrate the $NR$ of the matrix

$$\frac{1}{2}\begin{bmatrix} A + e^{2i\theta}\bar{A} & -i(A - e^{2i\theta}\bar{A}) \\ i(A - e^{2i\theta}\bar{A}) & A + e^{2i\theta}\bar{A} \end{bmatrix}$$

(see Figure 2.2).

From this, (2.8) is evident.

**3. Numerical range on an indefinite inner product.** Properties of the numerical range of a matrix on an indefinite inner product space have been presented

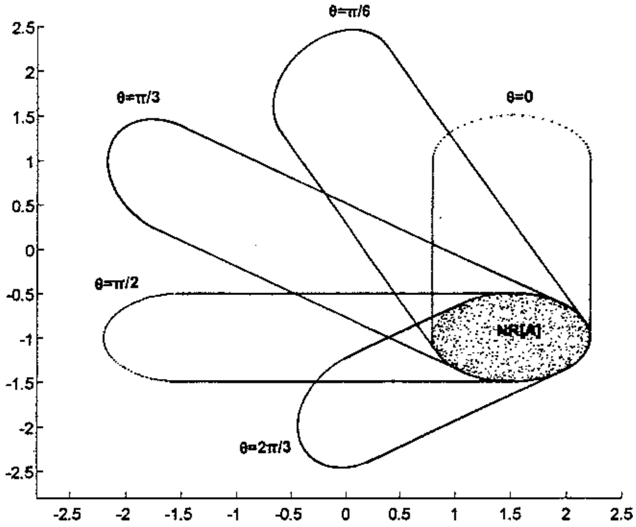


FIG. 2.2

in a recent publication [LTU]. The authors have extended some properties of the classical numerical range and have given a detailed description of $W_S^+(A)$ and

$$V_S^+(A) = \{\langle Ax, x\rangle : \quad x \in \mathbf{C}^n, \quad \langle x, x\rangle_S = 1\}.$$

Obviously, by (1.3), $W_S^+(A) = V_S^+(SA)$ and it is proved that $W_S^+(A)$ and $V_S^+(A)$ are always convex in a different approach from that in [B]. We next state a connection of $NR[A]$ and $W_S^+(A)$.

THEOREM 3.1. *For any hermitian matrix $S$, the set $NR[A] \cap W_S^+(A)$ is nonempty.*

*Proof.* For the hermitian $S$, let $P$ be an orthonormal matrix such that $S = P^*DP$, where $D$ is a real diagonal matrix. Then for any $S$-unit vector $x \in \mathbf{C}^n$ (i.e., $x^*Sx = 1$), we have

$$\langle Ax, x\rangle_S = x^*SAx = x^*P^*DPAx = y^*DBy = \langle By, y\rangle_D,$$

where $y = Px \in \mathbf{C}^n$ and $B = PAP^*$. Since $y$ is a $D$-unit vector, $W_S^+(A) = W_D^+(B)$. Moreover, for any unit vector $w \in \mathbf{C}^n$, the equations

$$w^*Aw = w^*P^*BPw = z^*Bz, \qquad z = Pw, \ \|z\| = 1$$

imply $NR[A] = NR[B]$.

Therefore, it is enough to prove that $NR[B] \cap W_D^+(B) \neq \emptyset$, where $D$ is a real diagonal matrix. Next we denote $D = \mathrm{diag}(D_1, -D_2)$, where $D_1 > 0$ and $D_2 \geq 0$ are the subdiagonal matrices. Let $\mathbf{E}$ be a subspace of $\mathbf{C}^n$ and $\xi_1, \ldots, \xi_k$ be a $D$-orthonormal basis of $\mathbf{E}$, i.e., $\langle \xi_j, \xi_i\rangle_D = \xi_i^* D\xi_j = \pm\delta_{ij}$, where $\delta_{ij}$ is Kronecker's symbol. In particular, we can consider that

$$1 = \langle \xi_i, \xi_i\rangle_D = \langle \xi_{io}, \xi_{io}\rangle_{D_1} \qquad \text{for } \xi_i = \begin{bmatrix} \xi_{io} \\ 0 \end{bmatrix}, \qquad i = 1, 2, \ldots, p$$

and

$$-1 = \langle \xi_j, \xi_j\rangle_D = \langle \xi_{jo}, \xi_{jo}\rangle_{D_2} \quad \text{for } \xi_j = \begin{bmatrix} 0 \\ \xi_{jo} \end{bmatrix}, \quad j = p+1, \ldots, k.$$

Then, for $\lambda \in W_D^+(B)$, there exists a vector $x_0 = [\xi_1, \dots, \xi_k]\,\eta$ of $\mathbf{E}$ such that

$$\lambda = \langle Bx_0, x_0\rangle_D = x_0^* D B x_0 = \eta^* \begin{bmatrix} \langle B\xi_1, \xi_1\rangle_D & \dots & \langle B\xi_k, \xi_1\rangle_D \\ \vdots & & \vdots \\ \langle B\xi_1, \xi_k\rangle_D & \dots & \langle B\xi_k, \xi_k\rangle_D \end{bmatrix} \eta$$

and

$$1 = \langle x_0, x_0\rangle_D = \eta^* \begin{bmatrix} \xi_1^* \\ \vdots \\ \xi_k^* \end{bmatrix} D\,[\xi_1 \dots \xi_k]\,\eta = \eta^* \mathrm{diag}\,(I_p, -I_{k-p})\,\eta.$$

For $\eta = [\begin{smallmatrix}\eta_1 \\ 0\end{smallmatrix}]$, with $\eta_1 \in \mathbf{C}^p$, clearly $\|\eta_1\| = 1$ and

$$\lambda = \eta_1^* \begin{bmatrix} \langle B\xi_1, \xi_1\rangle_D & \dots & \langle B\xi_p, \xi_1\rangle_D \\ \vdots & & \vdots \\ \langle B\xi_1, \xi_p\rangle_D & \dots & \langle B\xi_p, \xi_p\rangle_D \end{bmatrix} \eta_1$$

$$= \eta_1^* \begin{bmatrix} \langle B_1\xi_{10}, \xi_{10}\rangle_{D_1} & \dots & \langle B_1\xi_{p0}, \xi_{10}\rangle_{D_1} \\ \vdots & & \vdots \\ \langle B_1\xi_{10}, \xi_{p0}\rangle_{D_1} & \dots & \langle B_1\xi_{p0}, \xi_{p0}\rangle_{D_1} \end{bmatrix} \eta_1,$$

where $B_1$ is the $p \times p$ principal submatrix of $B$. Since $\xi_{10}, \dots, \xi_{p0}$ are $D_1$-orthonormal vectors and $D_1 > 0$, the vectors $D_1^{1/2}\xi_{10}, \dots, D_1^{1/2}\xi_{p0}$ are orthonormal, and by Theorem 2.1,

$$\lambda \in NR\left[D_1^{1/2} B_1 D_1^{-1/2}\right] = NR[B_1] \subset NR[B]. \qquad \square$$

Due to the fact that $W_{-S}^+(A) = W_{-D}^+(B)$ and $W_S(A) = W_S^+(A) \cup W_{-S}^+(A) = W_D^+(B) \cup W_{-D}^+(B)$, clearly we have the following.

COROLLARY 3.2. *For any hermitian matrix $S$, the sets $NR[A] \cap W_{-S}^+(A)$ and $NR[A] \cap W_S(A)$ are nonempty.*

COROLLARY 3.3. *For any hermitian matrix $S, NR[A] \cap V_S^+(A) \neq \emptyset$.*

*Proof.* Let $S = P^*DP$, where $D = \mathrm{diag}(D_1, -D_2)$ is a real diagonal matrix with $D_1 > 0$ and $D_2 \geq 0$. If $x$ is an $S$-unit vector, then the vector $y = Px$ is $D$-unit and $V_S^+(A) = V_D^+(B)$, where $A = P^*BP$. For $y = [\begin{smallmatrix}\theta \\ 0\end{smallmatrix}]$, we have

$$1 = \langle Dy, y\rangle = \theta^* D_1 \theta = (\theta^* D_1^{1/2})(D_1^{1/2}\theta)$$

and

$$\langle By, y\rangle = \langle B_1\theta, \theta\rangle = \langle \tilde{B}_1\theta_1, \theta_1\rangle,$$

where $B_1$ is a submatrix of $B$ and $\theta_1 = D_1^{1/2}\theta,\ B_1 = D^{1/2}\tilde{B}_1 D^{1/2}$.
Therefore,

$$NR[\tilde{B}_1] = NR[B_1] \subset NR[B] = NR[A]$$

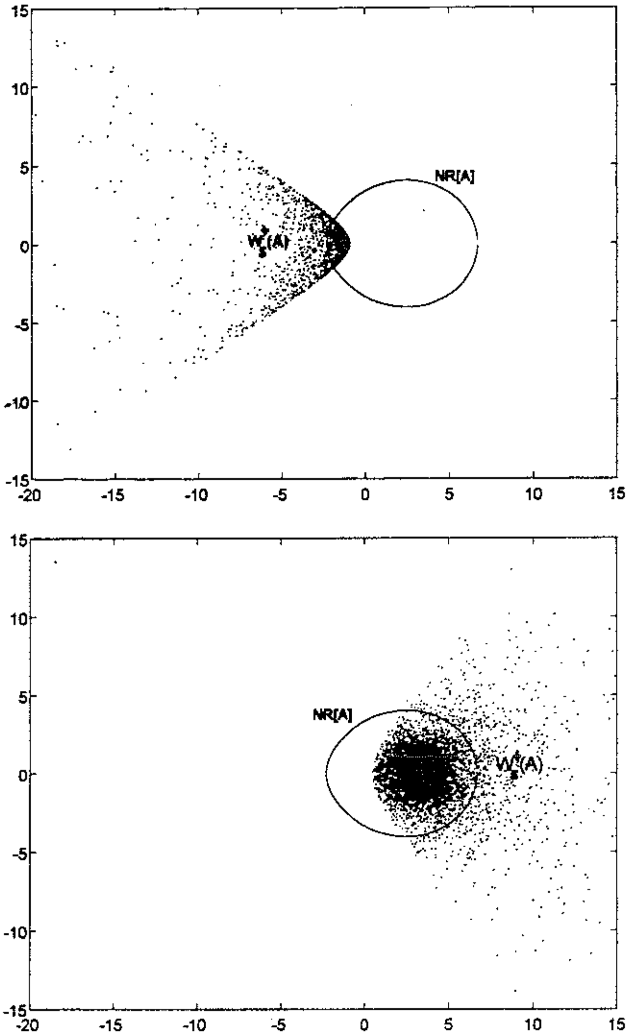


FIG. 3.1

and

$$NR[\tilde{B}_1] = V_{D_1}^{+}(B_1) \subset V_D^{+}(B) = V_S^{+}(A). \qquad \square$$

*Example* 3. The statement of Theorem 3.1 is illustrated in Figure 3.1 for

$$A = \begin{bmatrix} -1 & 0 & 2 \\ 0 & 2 & 8 \\ 1 & 0 & 3 \end{bmatrix},$$

first for $S = \operatorname{diag}(2, 0, -4)$, and then for $S = \operatorname{diag}(-2, 1, 4)$.

Another property of the numerical range of a matrix on an indefinite inner product, an analogue of Theorem 2.2, is as follows.

THEOREM 3.4. *For any matrix A,*

$$(3.1) W^+_{I_2\otimes S}\left(\begin{bmatrix} A & O \\ O & e^{2i\theta}\bar{A}\end{bmatrix}\right) = W^+_{I_2\otimes S}\left(\frac{1}{2}\begin{bmatrix} A+e^{2i\theta}\bar{A} & -i(A-e^{2i\theta}\bar{A}) \\ i(A-e^{2i\theta}\bar{A}) & A+e^{2i\theta}\bar{A}\end{bmatrix}\right),$$

*where* $0 \le \theta \le \pi$.

*Proof.* Let a vector $x \in \mathbf{C}^{2n}$ such that $x^*(I_2 \otimes S)x = 1$. Then

$$x^*(I_2\otimes S)\begin{bmatrix} A & O \\ O & e^{2i\theta}\bar{A}\end{bmatrix}x = \frac{1}{2}y^*(I_2\otimes S)\begin{bmatrix} A+e^{2iy}\bar{A} & -i(A-e^{2i\theta}\bar{A}) \\ i(A-e^{2i\theta}\bar{A}) & A+e^{2i\theta}\bar{A}\end{bmatrix}y,$$

where $y = \frac{1}{\sqrt{2}}\left(\begin{bmatrix} 1 & i \\ i & 1\end{bmatrix}\otimes I\right)x$. Since

$$y^*\,(I_2\otimes S)\,y \;=\; x^*\,(I_2\otimes S)\,x \;=\; 1,$$

(3.1) is obtained. □

On the other hand we have the following theorem.

THEOREM 3.5. *Let the hermitian matrix S have at least one positive eigenvalue. Then*

$$(3.2) \qquad W^+_{I_2\otimes S}\left(\begin{bmatrix} A & O \\ O & B\end{bmatrix}\right) = \text{Conv.hull}(W^+_S(A)\cup W^+_S(B)).$$

*Proof.* Let

$$\lambda \in W^+_{I_2\otimes S}\left(\begin{bmatrix} A & O \\ O & B\end{bmatrix}\right)$$

correspond to the $I_2 \otimes S$-unit vector $v = \left[\begin{smallmatrix} x \\ y\end{smallmatrix}\right]$ (i.e., $x^*Sx + y^*Sy = 1$). According to Lemma 2.2 in [LTU], we can consider that $x^*Sx > 0$ and $y^*Sy > 0$ and then

$$\lambda = v^*(I_2\otimes S)\begin{bmatrix} A & O \\ O & B\end{bmatrix}v = x^*SAx + y^*SBy = x^*Sx\frac{x^*SAx}{x^*Sx} + y^*Sy\frac{y^*SBy}{y^*Sy}.$$

Hence, $\lambda$ is a convex combination of

$$\frac{x^*SAx}{x^*Sx} \in W^+_S(A) \qquad \text{and} \qquad \frac{y^*SBy}{y^*Sy} \in W^+_S(B)$$

and consequently

$$W^+_{I_2\otimes S}\left(\begin{bmatrix} A & O \\ O & B\end{bmatrix}\right) \subset \text{Conv.hull}(W^+_S(A)\cup W^+_S(B)).$$

For the reverse relationship let $x^*Sx = 1$. Then the vector $z = \left[\begin{smallmatrix} x \\ 0\end{smallmatrix}\right]$ is $I_2 \otimes S$-unit and due to

$$z^*(I_2\otimes S)\begin{bmatrix} A & O \\ O & B\end{bmatrix}z = x^*SAx,$$

it is implied that

$$W_S^+(A) \subset W_{I_2\otimes S}^+\left(\begin{bmatrix} A & O \\ O & B \end{bmatrix}\right).$$

Similarly, for $y^*Sy = 1$,

$$W_S^+(B) \subset W_{I_2\otimes S}^+\left(\begin{bmatrix} A & O \\ O & B \end{bmatrix}\right).$$

Therefore,

$$W_S^+(A) \cup W_S^+(B) \subset W_{I_2\otimes S}^+\left(\begin{bmatrix} A & O \\ O & B \end{bmatrix}\right),$$

and due to the convexity of

$$W_{I_2\otimes S}^+\left(\begin{bmatrix} A & O \\ O & B \end{bmatrix}\right)$$

(3.2) follows. □

These two statements lead to an analogue of (2.2) in Theorem 2.2:

$$\begin{aligned} &\text{Conv.hull}(W_S^+(A) \cup W_S^+(e^{2i\theta}\bar{A})) \\ (3.3) \qquad &= W_{I_2\otimes S}^+\left(\frac{1}{2}\begin{bmatrix} A + e^{2i\theta}\bar{A} & -i(A - e^{2i\theta}\bar{A}) \\ i(A - e^{2i\theta}\bar{A}) & A + e^{2i\theta}\bar{A} \end{bmatrix}\right). \end{aligned}$$

Denoting $A = M + iN$ for $\theta = 0$, by (3.3) we obtain

$$\text{(3.4)} \qquad \text{Conv.hull}(W_S^+(A) \cup W_S^+(\bar{A})) = W_{I_2\otimes S}^+\left(\begin{bmatrix} M & N \\ -N & M \end{bmatrix}\right),$$

and by (3.1)

$$W_{I_2\otimes S}^+\left(\begin{bmatrix} A & O \\ O & \bar{A} \end{bmatrix}\right) = W_{I_2\otimes S}^+\left(\begin{bmatrix} M & N \\ -N & M \end{bmatrix}\right).$$

Similar results of (3.3) and (3.4) are obtained for $V_S^+(A)$.

**4. Reduction of *NR* of matrix polynomials.** Transferring the idea of Theorem 2.1 for the numerical range of matrix polynomials $L(\lambda)$ we state the next proposition.

THEOREM 4.1. *For $k \leq n$,*

$$\text{(4.1)} \qquad NR[L(\lambda)] = \bigcup_{\xi_1,\ldots,\xi_k} NR[M_\xi(\lambda)],$$

*where the $k \times k$ matrix polynomial*

$$M_\xi(\lambda) = G^*(E \otimes L(\lambda))G,$$

*with*

$$E = \begin{bmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{bmatrix}_{k\times k}, \qquad G = \operatorname{diag}(\xi_1, \dots, \xi_k)_{nk\times k},$$

*and $\xi_1, \dots, \xi_k$ are orthonormal vectors of $\mathbf{C}^n$. In (4.1) the union runs over all sets by $k$ orthonormal vectors of $\mathbf{C}^n$.*

*Proof.* Let the vectors $\xi_1, \dots, \xi_k \in \mathbf{C}^n$ be an orthonormal basis of a subspace $\mathbf{E} \subset \mathbf{C}^n$ and let $x \in \mathbf{E}$. Then

$$x = [\xi_1 \dots \xi_k]y, \quad y \in \mathbf{C}^k$$

and

$$\begin{aligned} x^*L(\lambda)x &= y^*[\xi_1 \dots \xi_k]^* L(\lambda)[\xi_1 \dots \xi_k]y \\ &= y^* \begin{bmatrix} \xi_1^* L(\lambda)\xi_1 & \dots & \xi_1^* L(\lambda)\xi_k \\ \vdots & & \vdots \\ \xi_k^* L(\lambda)\xi_1 & \dots & \xi_k^* L(\lambda)\xi_k \end{bmatrix} y = y^* M_\xi(\lambda) y. \end{aligned}$$

Hence, running over all sets by $k$ orthonormal vectors, (4.1) results. □

Obviously (4.1) for $k = 2$ is simple. Denoting $L(\lambda)$ as

$$\overline{L}(\lambda) = I\lambda^q + \bar{A}_1\lambda^{q-1} + \dots + \bar{A}_q,$$

a nearly similar relation to (2.2) arises.

THEOREM 4.2.

$$\begin{aligned} &NR\left(\begin{bmatrix} L(\lambda) & O \\ O & e^{2i\theta}\overline{L}(\lambda) \end{bmatrix}\right) \\ (4.2) \qquad &= NR\left(\frac{1}{2}\begin{bmatrix} L(\lambda) + e^{2i\theta}\overline{L}(\lambda) & -i(L(\lambda) - e^{2i\theta}\overline{L}(\lambda)) \\ i(L(\lambda) - e^{2i\theta}\overline{L}(\lambda)) & L(\lambda) + e^{2i\theta}\overline{L}(\lambda) \end{bmatrix}\right). \end{aligned}$$

*Proof.* Let the vector $z \in \mathbf{C}^{2n}$. Then, as in Theorem 2.2,

$$z^* \begin{bmatrix} L(\lambda) & O \\ O & e^{2i\theta}\overline{L}(\lambda) \end{bmatrix} z = \frac{1}{2} w^* \begin{bmatrix} L(\lambda) + e^{2i\theta}\overline{L}(\lambda) & -i(L(\lambda) - e^{2i\theta}\overline{L}(\lambda)) \\ i(L(\lambda) - e^{2i\theta}\overline{L}(\lambda)) & L(\lambda) + e^{2i\theta}\overline{L}(\lambda) \end{bmatrix} w,$$

where

$$w = \frac{1}{\sqrt{2}} \begin{bmatrix} I & iI \\ iI & I \end{bmatrix} z.$$

Note that $\|w\| = \|z\|$. □

Writing $L(\lambda) = L_1(\lambda) + iL_2(\lambda)$, where the coefficients of $L_1(\lambda)$ and $L_2(\lambda)$ are real matrices, by (4.2) and for $\theta = 0$, we have the following corollary.

COROLLARY 4.3.

$$NR\left(\begin{bmatrix} L(\lambda) & O \\ O & \overline{L}(\lambda) \end{bmatrix}\right) = NR\left(\begin{bmatrix} L_1(\lambda) & L_2(\lambda) \\ -L_2(\lambda) & L_1(\lambda) \end{bmatrix}\right).$$

REFERENCES

[AL] T. Ando and C. K. Li, *The numerical range and numerical radii*, Linear and Multilinear Algebra, 37 (1994), pp. 1–3.

[B] T. Bayasgalan, *The numerical range of linear operators in spaces with an indefinite metric*, Acta Math. Hungar., 57 (1991), pp. 7–9.

[GLR] I. Gohberg, P. Lancaster, and L. Rodman, *Matrices and Indefinite Scalar Products*, Oper. Theory Adv. Appl. 8, Birkhauser, Basel, 1983.

[HJ] R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, London. 1991.

[LTU] C.-K. Li, N.-K. Tsing, and F. Uhlig, *Numerical ranges of an indefinite inner product space*, Electron. J. Linear Algebra, 1 (1996), pp. 1–17.

[LMZ] P. Lancaster, J. Maroulas, and P. Zizler, *The Numerical Range of Selfadjoint Matrix Polynomials*, Oper. Theory Adv. Appl. 106, Birkhauser, Basel, 1998, pp. 291–304.

[LR] C.-K. Li and L. Rodman, *Numerical range of matrix polynomials*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 1256–1265.

[M] A. Markus, *Introduction to Spectral Theory of Polynomial Operator Pencils*, American Mathematical Society, Providence, RI, 1988.

[MMP] A. Markus, J. Maroulas, and P. Psarrakos, *Spectral Properties of a Matrix Polynomial Connected with a Component of its Numerical Range*, Oper. Theory Adv. Appl. 106, Birkhauser, Basel, 1998, pp. 305–308.

[MP] M. Markus and C. Pesce, *Computer generated numerical ranges and some resulting theorems*, Linear and Multilinear Algebra, 20 (1987), pp. 121–157.

[MP1] J. Maroulas and P. Psarrakos, *The boundary of numerical range of matrix polynomials*, Linear Algebra Appl., 267 (1997), pp. 101–111.

[MP2] J. Maroulas and P. Psarrakos, *On the connectedness of numerical range of matrix polynomials*, Linear Algebra Appl., 280 (1998), pp. 97–108.

[MP3] J. Maroulas and P. Psarrakos, *A connection between numerical ranges of selfadjoint matrix polynomials*, Linear and Multilinear Algebra, 44 (1998), pp. 327–340.