The Electronic Journal of Linear Algebra.
A publication of the International Linear Algebra Society.
Volume 6, pp. 20-30, March 2000.
ISSN 1081-3810. http://math.technion.ac.il/iic/ela

ELA

ON THE RELATION BETWEEN THE NUMERICAL RANGE AND THE JOINT NUMERICAL RANGE OF MATRIX POLYNOMIALS*

P. J. PSARRAKOS† AND M. J. TSATSOMEROS‡

Abstract. It is shown that the numerical range, $\mathrm{NR}[P(\lambda)]$, of a matrix polynomial $P(\lambda) = A_m\lambda^m + \ldots + A_1\lambda + A_0$ consists of the roots of all scalar polynomials whose coefficients correspond to the elements of the convex hull of the joint numerical range of the $(m+1)$-tuple $(A_0, A_1, \ldots, A_m)$. Moreover, the elements of the joint numerical range that give rise to scalar polynomials with a common root belonging to $\mathrm{NR}[P(\lambda)]$ form a connected set. The latter fact is used to examine the multiplicity of roots belonging to the intersection of the root zones of $\mathrm{NR}[P(\lambda)]$. Also an approximation scheme for $\mathrm{NR}[P(\lambda)]$ is proposed, in terms of numerical ranges of diagonal matrix polynomials.

Key words. joint numerical range, matrix polynomial, root zone

AMS subject classifications. 15A60, 47A12

1. Introduction.

The numerical range of a matrix polynomial and the joint numerical range of a set of matrices (usually taken to be the coefficients of a matrix polynomial) are both notions that generalize the classical numerical range of a matrix. The former is a one-dimensional generalization and the latter a multi-dimensional generalization. We will continue previous efforts found in our references to study the relations between these sets and discover their properties. The concluding paragraph of this section includes a brief description of our results, requiring the following basic terminology and notation.

Let $\mathcal{M}_n(\mathbb{C})$ denote the algebra of $n \times n$ matrices over the complex field and let

$$P(\lambda) = A_m\lambda^m + \ldots + A_1\lambda + A_0 \tag{1}$$

be a *matrix polynomial*, where $A_j \in \mathcal{M}_n(\mathbb{C})$ $(j = 0, 1, \ldots, m)$, $A_m \neq 0$. The positive integers m and n are referred to as the *degree* and the *order* of $P(\lambda)$, respectively. The scalar $\lambda_0 \in \mathbb{C}$ is called an *eigenvalue* of $P(\lambda)$ if the equation $P(\lambda_0)x = 0$ has a nonzero solution $x = x_0 \in \mathbb{C}^n$. Thus the *spectrum* of $P(\lambda)$ is the set

$$\sigma[P(\lambda)] = \{\lambda \in \mathbb{C} \ : \ \det P(\lambda) = 0\}.$$

As $\det P(\lambda)$ is a polynomial of degree at most $n \cdot m$, $\sigma[P(\lambda)]$ is either equal to $\mathbb{C}$ or it consists of at most $n \cdot m$ complex numbers. The *numerical range* of $P(\lambda)$ is the set

$$\mathrm{NR}[P(\lambda)] = \{\lambda \in \mathbb{C} \ : \ x^*P(\lambda)x = 0 \ \text{ for some nonzero } x \in \mathbb{C}^n\},$$

*Received by the editors on 6 May 1999. Accepted for publication on 27 February 2000. Handling editor: Bryan L. Shader

†Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan, Canada S4S 0A2 (panos@math.uregina.ca).

‡Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan, Canada S4S 0A2 (tsat@math.uregina.ca). Research supported by NSERC research grant.

which is closed and contains $\sigma[P(\lambda)]$. Notice that the notion of $\mathrm{NR}[P(\lambda)]$ generalizes the classical numerical range of a matrix A, that is, the convex set

$$\mathrm{NR}[\lambda I - A] = \{x^* A x \ : \ x \in \mathbb{C}^n, x^* x = 1\}.$$

Also notice that $\lambda_0 \in \mathrm{NR}[P(\lambda)]$ if and only if $0 \in \mathrm{NR}[\lambda I - P(\lambda_0)]$. It is known that $\mathrm{NR}[P(\lambda)]$ is not necessarily connected and that it is bounded if and only if $0 \notin \mathrm{NR}[\lambda I - A_m]$; see Li and Rodman [6].

The *joint numerical range* of $P(\lambda)$ is the set of complex $(m+1)$-tuples

$$\mathrm{JNR}[P(\lambda)] = \{(x^* A_0 x, x^* A_1 x, \dots, x^* A_m x) \ : \ x \in \mathbb{C}^n, \ x^* x = 1\}.$$

The joint numerical range, being a continuous image of the unit sphere, is compact and connected but not necessarily convex; see Binding and Li [1]. Its convex hull is denoted by $\mathrm{Co}\{\mathrm{JNR}[P(\lambda)]\}$. Observe that

$$\mathrm{NR}[P(\lambda)] = \{\lambda \in \mathbb{C} : c_m \lambda^m + \ldots + c_1 \lambda + c_0 = 0, \ \ (c_0, c_1, \dots, c_m) \in \mathrm{JNR}[P(\lambda)]\}.$$

Consider now the following two subsets of $\mathbb{C}$: the set of roots of all m-degree scalar polynomials whose coefficients correspond to elements of the joint numerical range of $m+1$ given matrices, and the set of roots of all m-degree scalar polynomials whose coefficients correspond to elements of the convex hull of the joint numerical range of the same $m+1$ matrices. As we will see, both these sets coincide and are equal to the numerical range of the associated matrix polynomial, despite the fact that the joint numerical range is not necessarily convex (Proposition 2.1). Moreover, the elements of the joint numerical range associated with scalar polynomials that have a common root belonging to the numerical range, form a set that is not only nonempty but also connected (Proposition 2.2). Whether or not the multiplicities of roots of scalar polynomials with coefficients from the two sets above also carry over is an open (and seemingly substantial) problem with implications in the factorization of matrix polynomials. Some partial results in this regard can be found in section 3. Finally, in section 4 we describe a process to approximate the numerical range of a matrix polynomial using the (joint) numerical ranges of diagonal matrices.

2. The convex hull of the joint numerical range. Notice that if $Q(\lambda)$ is a matrix polynomial of degree m, then

$$\mathrm{JNR}[P(\lambda)] \subseteq \mathrm{JNR}[Q(\lambda)] \implies \mathrm{NR}[P(\lambda)] \subseteq \mathrm{NR}[Q(\lambda)].$$

More generally, given a set G of complex $(m+1)$-tuples such that $\mathrm{JNR}[P(\lambda)] \subseteq G$,

$$\mathrm{NR}[P(\lambda)] \ \subseteq \{\lambda \in \mathbb{C} : b_m \lambda^m + \ldots + b_1 \lambda + b_0 = 0, \ \ (b_0, b_1, \dots, b_m) \in G\}. \tag{2}$$

The following strengthening of (2) holds when $G = \mathrm{Co}\{\mathrm{JNR}[P(\lambda)]\}$.

PROPOSITION 2.1. *For every matrix polynomial $P(\lambda)$ of degree m we have*

$$\mathrm{NR}[P(\lambda)] = \{\lambda \in \mathbb{C} : c_m \lambda^m + \ldots + c_1 \lambda + c_0 = 0, \ (c_0, c_1, \dots, c_m) \in \mathrm{Co}\{\mathrm{JNR}[P(\lambda)]\}\}.$$

Proof. Let $\lambda_0 \in \mathbb{C}$ such that

$$c_m\lambda_0^m + \ldots + c_1\lambda_0 + c_0 = 0 \tag{3}$$

for some $\mathbf{c} = (c_0, c_1, \ldots, c_m) \in \mathrm{Co}\{\mathrm{JNR}[P(\lambda)]\}$. By Caratheodory's Theorem applied to $\mathbb{C}^{m+1}$, $\mathbf{c}$ is a convex combination of $\mu \le 2(m+1)+1$ points $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_\mu \in \mathrm{JNR}[P(\lambda)]$, that is,

$$\mathbf{c} = \sum_{j=1}^{\mu} t_j \mathbf{q}_j, \quad \sum_{j=1}^{\mu} t_j = 1, \quad t_j \ge 0 \quad (j = 1, 2, \ldots, \mu).$$

Since there exist unit vectors $x_j \in \mathbb{C}^n$ such that

$$\mathbf{q}_j = (x_j^* A_0 x_j, x_j^* A_1 x_j, \ldots, x_j^* A_m x_j) \quad (j = 1, 2, \ldots, \mu),$$

we have that

$$c_k = \sum_{j=1}^{\mu} t_j (x_j^* A_k x_j) \quad (k = 1, 2, \ldots, m)$$

and thus (3) becomes

$$\sum_{j=1}^{\mu} t_j \left((x_j^* A_m x_j)\lambda_0^m + \ldots + (x_j^* A_1 x_j)\lambda_0 + (x_j^* A_0 x_j) \right) = \sum_{j=1}^{\mu} t_j \left(x_j^* P(\lambda_0) x_j \right) = 0.$$

Since all the complex numbers $x_j^* P(\lambda_0) x_j \;\; (j = 1, 2, \ldots, \mu)$ belong to the convex set $\mathrm{NR}[\lambda I - P(\lambda_0)]$, it follows that $0 \in \mathrm{NR}[\lambda I - P(\lambda_0)]$, i.e., $\lambda_0 \in \mathrm{NR}[P(\lambda)]$. Hence

$$\{\lambda \in \mathbb{C} : c_m\lambda^m + \ldots + c_1\lambda + c_0 = 0, \;\; (c_0, c_1, \ldots, c_m) \in \mathrm{Co}\{\mathrm{JNR}[P(\lambda)]\} \subseteq \mathrm{NR}[P(\lambda)].$$

The opposite inclusion follows from the discussion preceding this proposition. ☐

We continue with an interesting geometric feature of the relation between $\mathrm{NR}[P(\lambda)]$ and $\mathrm{JNR}[P(\lambda)]$. Suppose that $c_m\lambda_0^m + \ldots + c_1\lambda_0 + c_0 = 0$. Then there exist $a_0, a_1, \ldots, a_{m-1} \in \mathbb{C}$ such that

$$\begin{aligned} c_m\lambda^m + \ldots + c_1\lambda + c_0 &= (\lambda - \lambda_0)(a_{m-1}\lambda^{m-1} + \ldots + a_1\lambda + a_0) \\ &= a_{m-1}\lambda^m + (a_{m-2} - \lambda_0 a_{m-1})\lambda^{m-1} + \ldots + (a_0 - \lambda_0 a_1)\lambda - \lambda_0 a_0. \end{aligned}$$

We may then define the set of complex $(m+1)$-tuples

$$\begin{aligned} T_1(\lambda_0) &= \{(c_0, c_1, \ldots, c_m) : c_m\lambda_0^m + \ldots + c_1\lambda_0 + c_0 = 0\} \\ &= \{(-\lambda_0 a_0, a_0 - \lambda_0 a_1, \ldots, a_{m-1}) : (a_0, a_1, \ldots, a_{m-1}) \in \mathbb{C}^m\} \\ &= \mathrm{span}\{(-\lambda_0, 1, 0, \ldots, 0), (0, -\lambda_0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, -\lambda_0, 1)\}. \end{aligned} \tag{4}$$

By Proposition 2.1, if $T_1(\lambda_0)$ intersects $\mathrm{Co}\{\mathrm{JNR}[P(\lambda)]\}$, it also intersects $\mathrm{JNR}[P(\lambda)]$. However, the following more general result holds.

Proposition 2.2. *Let $P(\lambda) = A_m\lambda^m + \ldots + A_1\lambda + A_0$ be an $n \times n$ matrix polynomial with $n \geq 3$. For every $\omega \in \mathrm{NR}[P(\lambda)]$ the set $T_1(\omega) \cap \mathrm{JNR}[P(\lambda)]$ is connected.*

Proof. The set $T_1(\omega) \cap \mathrm{JNR}[P(\lambda)]$ consists of all points $(c_0, c_1, \ldots, c_m)$ such that $c_m\omega^m + \ldots + c_1\omega + c_0 = 0$ and $(c_0, c_1, \ldots, c_m) = (x_0^* A_0 x_0, x_0^* A_1 x_0, \ldots, x_0^* A_m x_0)$ for some $x_0 \in \mathbb{C}^n$. It is shown in Lyubich and Markus [8] that the set

$$L_0 = \{x \in \mathbb{C}^n : x^*x = 1 \text{ and } x^*P(\omega)x = 0\}$$

is connected for $n \geq 3$. Consequently, if we consider the continuous function

$$\Phi : \{x \in \mathbb{C}^n : x^*x = 1\} \longrightarrow \mathrm{JNR}[P(\lambda)]$$

defined by $\Phi(x) = (x^*A_0x, x^*A_1x, \ldots, x^*A_mx)$, then $\Phi(L_0) = T_1(\omega) \cap \mathrm{JNR}[P(\lambda)]$ is indeed a connected set. □

In the special case when $P(\lambda)$ is self-adjoint, that is, the coefficient matrices are Hermitian, $\mathrm{JNR}[P(\lambda)]$ consists of real $(m+1)$-tuples. Therefore for every $\omega \in \mathrm{NR}[P(\lambda)] \cap \mathbb{R}$ we can define the real hyperplane

$$\begin{aligned} T_1(\omega)^{\mathbb{R}} &= \{(-\omega a_0, \ldots, a_{m-2} - \omega a_{m-1}, a_{m-1}) : (a_0, a_1, \ldots, a_{m-1}) \in \mathbb{R}^m\} \\ &= \mathrm{span}\{(-\omega, 1, 0, \ldots, 0), (0, -\omega, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, -\omega, 1)\}. \end{aligned}$$

Similarly to Proposition 2.2, it can be shown that if $P(\lambda)$ is an $n \times n$ matrix polynomial with $n \geq 3$, then $T_1(\omega)^{\mathbb{R}} \cap \mathrm{JNR}[P(\lambda)]$ is connected for every $\omega \in \mathrm{NR}[P(\lambda)] \cap \mathbb{R}$.

3. Root zones of the numerical range. Let $P(\lambda)$ be a matrix polynomial as in (1), whose numerical range is bounded, that is, $0 \notin \mathrm{NR}[\lambda I - A_m]$. We can always number the roots of $x^*P(\lambda)x = 0$ to obtain a sequence of functionals $\lambda_1(x), \lambda_2(x), \ldots, \lambda_m(x)$ on the unit sphere S, having ranges $\Lambda_1, \Lambda_2, \ldots, \Lambda_m$, respectively. We will refer to $\Lambda_1, \Lambda_2, \ldots, \Lambda_m$ as *root zones* of $\mathrm{NR}[P(\lambda)]$. We clearly have that $\mathrm{NR}[P(\lambda)] = \bigcup_{j=1}^m \Lambda_j$; see Isaev [5]. Lyubich [7] refers to the Λ_j as 'root branches' and uses them in identifying conditions that are equivalent to the roots of $x^*P(\lambda)x = 0$ being simple at every point of S.

If there exists a nonzero $x \in \mathbb{C}^n$ such that $x^*P(\lambda)x = 0$ has multiple roots, then some root zones have nonempty intersection and thus $\mathrm{NR}[P(\lambda)]$ has fewer than m connected components. In what follows, we will consider the converse, namely, whether the nonempty intersection of some of the root zones necessarily implies the existence of multiple roots; see Maroulas and Psarrakos [9, Theorem 3.1] for the self-adjoint case. For that purpose, we will use the notion of the joint numerical range and the results of the previous section. We will also need to restrict our attention to a compact subset of S, where the functionals $\lambda_1(x), \lambda_2(x), \ldots, \lambda_m(x)$ are well defined and continuous. We expand on this requirement next.

Certainly, the roots of $x^*P(\lambda)x = 0$ depend continuously on the coefficients x^*A_jx, and thus on x. This fact can be deduced from Ostrowski [10, pp. 334–335] by considering the companion matrix of the polynomial

$$\lambda^m + b_{m-1}\lambda^{m-1} + \ldots + b_1\lambda + b_0,$$

where

$$b_j = \frac{x^* A_j x}{x^* A_m x} \quad (j = 0, 1, \ldots, m-1).$$

In spite of the continuous dependence of the roots of $x^* P(\lambda) x = 0$ on x, it may not be possible to define the aforementioned functionals $\lambda_1(x), \lambda_2(x), \ldots, \lambda_m(x)$ continuously on the whole unit sphere S. This situation is illustrated by the following example. Consider the matrix polynomial

$$P(\lambda) = I\ \lambda^2 - \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}$$

and let $x = (\cos\theta e^{-i\phi}, \sin\theta)^T$ parameterize S. We then have that

$$x^* P(\lambda) x = \lambda^2 - \sin(2\theta) e^{i\phi}.$$

Thus, if we require continuity of the roots $\lambda_1(x), \lambda_2(x)$ as functions of x, then it is easy to verify that they can not be defined globally on S.

When the roots of $x^* P(\lambda) x = 0$ are simple for every $x \in S$, a unique global definition of each $\lambda_j(x)$ as a continuous functional on S is possible and is indeed obtained in Lyubich [7]. Otherwise, to proceed with our goal, we need to modify the notion of a root branch $\lambda_j(x)$, using in essence the arguments in [7, pp. 55–56]. We will define continuous functionals $\lambda_j = \lambda_j(x)$ on a certain compact subset M_* of S, using notions from differential geometry for which our general reference is Helgason [4]. We construct M_* under the assumption that the polynomial

$$L(\lambda, x) \equiv x^* P(\lambda) x \in \mathbb{C}[\lambda, x_1, x_2, \ldots, x_n]$$

has no multiple irreducible factors. Let $D(x)$ be the discriminant of $L(\lambda, x)$ with respect to the indeterminate λ, that is, $D(x)$ is the resultant of the polynomial $L(\lambda, x)$ and its derivative $\frac{\partial}{\partial\lambda} L(\lambda, x)$. If $D(x) \equiv 0$ in the polynomial ring $\mathbb{C}[\lambda, x_1, x_2 \ldots, x_n]$, then $L(\lambda, x)$ has a repeated factor not in $\mathbb{C}[x]$; cf. Walker [11, p. 25]). Therefore, under our assumption, $D(x)$ is a non-zero element of $\mathbb{C}[x]$.

Consider now the set

$$\begin{aligned} U &= [\, \mathbb{C}^n \setminus \{0\}\,] \setminus \{x \in \mathbb{C}^n \ :\ L(\lambda, x) = \frac{\partial}{\partial\lambda} L(\lambda, x) = 0 \ \text{ for some } \lambda \in \mathbb{C}\} \\ &= \{x \in \mathbb{C}^n \setminus \{0\} : D(x) \neq 0\}. \end{aligned}$$

The set U is open and is the complement of an algebraic hypersurface in $\mathbb{C}^n$. A complex algebraic hypersurface has real codimension 2 in the whole space. Thus by [4, p. 346, Corollary 12.5], U is arcwise connected. It is clear that U is dense in $\mathbb{C}^n$. Therefore the set

$$U_1 = \{x \in \mathbb{C}^n : x^* x = 1, D(x) \neq 0\}$$

is an arcwise connected dense open subset of the unit sphere S. The set U_1 is not necessarily simply connected. Consider the universal covering space M_0 of U_1. We can define on M_0 analytic functionals $\lambda_1, \lambda_2, \ldots, \lambda_m$ being roots of $x^*P(\lambda)x = 0$, via a method similar to that mentioned in [7]. Denote by π_0 the covering map of M_0 onto U_1. For every $p \in M_0$, these functionals satisfy $\lambda_i(p) \neq \lambda_j(p)$ for $1 \leq i < j \leq m$. For every pair p, q of points of M_0 satisfying $\pi_0(p) = \pi_0(q)$, there is a permutation σ of $\{1, 2, \ldots, m\}$ for which

$$\lambda_i(q) = \lambda_{\sigma(i)}(p) \quad (i = 1, 2, \ldots, m).$$

Define now an equivalence relation $\sim$ on M_0 as follows.

$$p \sim q \;\text{ if and only if }\; \pi_0(p) = \pi_0(q) \;\text{ and }\; \lambda_i(p) = \lambda_i(q) \quad (i = 1, 2, \ldots, m).$$

Define now M as the coset space $M_0/\sim$. Then M is also a covering space of U_1. Let π denote the covering map of M onto U_1. The functionals λ_j are well defined on M. For every pair p, q of M for which $\pi(p) = \pi(q)$ and $p \neq q$, there is an index $1 \leq i \leq m$ with $\lambda_i(p) \neq \lambda_i(q)$. For every $x \in U_1$, the number L of points $p \in M$ for which $\pi(p) = x$ is the same and not greater than $m!$. Thus, M is an analytic manifold with the natural Riemannian structure. Denote by d the metric function on M or S. As earlier, using the results in [10, pp. 334–335], we can in fact deduce that the functionals λ_j are uniformly continuous on M. Taking an arbitrary $\epsilon > 0$, as S is compact, there are points $y_1, y_2, \ldots, y_\ell$ in S such that

$$S = \bigcup_{j=1}^{\ell} \{y \in S \;:\; d(y, y_j) < \epsilon/2\}.$$

Since U_1 is dense in S, there exist $x_1, \ldots, x_\ell \in U_1$ with $d(x_j, y_j) < \epsilon/2$ $(j = 1, 2, \ldots, \ell)$. Thus

$$U_1 = \bigcup_{j=1}^{\ell} \{x \in U_1 \;:\; d(x, x_j) < \epsilon\}.$$

For each j, there exist points $p_{j,1}, p_{j,2}, \ldots, p_{j,L}$ of M for which $\pi(p_{j,k}) = x_j$ $(k = 1, 2, \ldots, L)$ and $p_{j,s} \neq p_{j,t}$ for $1 \leq s < t \leq L$. Then,

$$M = \bigcup_{j=1}^{\ell} \bigcup_{k=1}^{L} \{x \in M \;:\; d(x, p_{j,k}) < \epsilon\}.$$

Since $\epsilon > 0$ is arbitrary, the space M is totally bounded. The completion M_* of M is compact. By uniform continuity, the functionals λ_j are defined and continuous on M_*.

The map π can be extended to a continuous map of M_* onto the unit sphere S. Based on the above discussion and notation, we will in the sequel refer to the *root zones* of $\mathrm{NR}[P(\lambda)]$ (where $P(\lambda)$ is a matrix polynomial as in (1) with bounded numerical range) as the sets

$$\Lambda_j = \{\lambda_j(x) : x \in M_*\} \quad (j = 1, 2, \ldots, m).$$

Each root zone Λ_j is a compact subset of the Gaussian plane $\mathbb{C}$.

For $\lambda_0 \in \mathrm{NR}[P(\lambda)]$ consider now a compact set $F \subseteq M_*$ that contains a vector x_0 such that $x_0^* P(\lambda_0) x_0 = 0$. Viewing $\lambda_j(x)$ as a function of the coefficients of $x^* P(\lambda) x$, define the sets

$$V_{j,F}(\lambda_0) = \{(c_0, c_1, \ldots, c_m) = (x^* A_0 x, x^* A_1 x, \ldots, x^* A_m x) : x \in F \text{ and } \lambda_0 = \lambda_j(x) := \lambda_j(c_0, c_1, \ldots, c_m)\} \quad (j = 1, 2, \ldots, m)$$

and

$$T_F(\lambda_0) = \{(c_0, c_1, \ldots, c_m) = (x^* A_0 x, x^* A_1 x, \ldots, x^* A_m x) : x \in F\} \cap T_1(\lambda_0),$$

where $T_1(\lambda_0)$ is as defined in (4). Clearly,

$$T_F(\lambda_0) \;=\; \bigcup_{j=1}^{m} V_{j,F}(\lambda_0).$$

THEOREM 3.1. *Consider the $n \times n$ matrix polynomial $P(\lambda) = A_m \lambda^m + \ldots + A_1 \lambda + A_0$. Assume that $0 \notin \mathrm{NR}[\lambda I - A_m]$ and let Λ_i, Λ_j be distinct root zones of $\mathrm{NR}[P(\lambda)]$. If*

$$\omega \in (\Lambda_i \cap \Lambda_j) \setminus (\bigcup_{k \neq i,j} \Lambda_k)$$

and for a compact $F \subseteq M_*$ there exist $x_1, x_2 \in F$ such that $\lambda_i(x_1) = \lambda_j(x_2) = \omega$, and if the set $\{x \in S : x^* P(\omega) x = 0\} \cap F$ is connected, then there exists $x_0 \in F$ such that $\lambda_i(x_0) = \lambda_j(x_0)$, namely, ω is a double root of $x_0^* P(\lambda) x_0 = 0$.

Proof. Note that since the set $\{x \in S : x^* P(\omega) x = 0\} \cap F$ is connected, following the proof of Proposition 2.2, we can show that $T_F(\omega)$ is a connected subset of $\mathrm{JNR}[P(\lambda)]$. Moreover,

$$T_F(\omega) = V_{i,F}(\omega) \cup V_{j,F}(\omega),$$

where these three sets are compact subsets of $\mathrm{JNR}[P(\lambda)]$. Consequently, $V_{i,F}(\omega) \cap V_{j,F}(\omega) \neq \emptyset$ and there exists $x_0 \in F$ such that $\lambda_i(x_0) = \lambda_j(x_0) = \omega$. $\square$

We comment that when $\omega \in (\Lambda_i \cap \Lambda_j) \setminus (\bigcup_{k \neq i,j} \Lambda_k)$, then there always exist $p, q \in M_*$ such that $\lambda_i(p) = \lambda_j(q) = \omega$. However, for the existence of the vector x_0 as in the above theorem, p and q must be connected by a continuous curve in

$$\{x \in S : x^* P(\omega) x = 0\} \cap M_*,$$

a fact implied by the remaining assumptions of Theorem 3.1.

THEOREM 3.2. *Consider the $n \times n$ matrix polynomial $P(\lambda) = A_m \lambda^m + \ldots + A_1 \lambda + A_0$. Assume that $0 \notin \mathrm{NR}[\lambda I - A_m]$ and let Λ_i, Λ_j be distinct root zones of $\mathrm{NR}[P(\lambda)]$. If there exists a compact $F \subseteq M_*$ and $x_1, x_2 \in F$ such that $\lambda_i(x_1) = \lambda_j(x_2) = \omega$, and if the set $\{x \in S : x^* P(\omega) x = 0\} \cap F$ is connected, then there exists $x_0 \in F$ such that ω is a multiple root of $x_0^* P(\lambda) x_0 = 0$.*

ELA

Proof. As in the proof of Theorem 3.1, the sets

$$V_{i,F}(\omega) \quad \text{and} \quad \bigcup_{k\neq i} V_{k,F}(\omega)$$

are nonempty and compact subsets of $F \subseteq M_*$. Moreover, the set

$$T_F(\omega) = V_{i,F}(\omega) \cup \left[\bigcup_{k\neq i} V_{k,F}(\omega) \right]$$

is a compact and connected subset of $\mathrm{JNR}[P(\lambda)]$. Hence,

$$V_{i,F}(\omega) \cap \left[\bigcup_{k\neq i} V_{k,F}(\omega) \right] \neq \emptyset,$$

that is, there exists $x_0 \in F$ such that $\lambda_i(x_0) = \lambda_k(x_0) = \omega$ for at least one $k \neq i$. ☐

4. Polyhedra and the joint numerical range. Consider two matrix polynomials $P(\lambda) = A_m\lambda^m + \ldots + A_1\lambda + A_0$ and $Q(\lambda) = B_m\lambda^m + \ldots + B_1\lambda + B_0$, not necessarily of the same order. If $\mathrm{JNR}[P(\lambda)] \subseteq \mathrm{JNR}[Q(\lambda)]$ or if $\mathrm{Co}\{\mathrm{JNR}[P(\lambda)]\} \subseteq \mathrm{Co}\{\mathrm{JNR}[Q(\lambda)]\}$, then by our discussion in section 2, $\mathrm{NR}[P(\lambda)] \subseteq \mathrm{NR}[Q(\lambda)]$. Let us further assume $\mathrm{NR}[P(\lambda)]$ and $\mathrm{NR}[Q(\lambda)]$ are bounded. Then for any $(c_0, c_1 \ldots, c_m)$ in $\mathrm{JNR}[P(\lambda)]$, the equation $c_m\lambda^m + \ldots + c_1\lambda + c_0 = 0$ has at least one root in every connected component of $\mathrm{NR}[P(\lambda)]$. A similar statement is true for $\mathrm{JNR}[Q(\lambda)]$ and the connected components of $\mathrm{NR}[Q(\lambda)]$. Thus $\mathrm{NR}[Q(\lambda)]$ has at most as many connected components as $\mathrm{NR}[P(\lambda)]$.

In view of the above, consider now two diagonal matrix polynomials $D_1(\lambda)$ and $D_2(\lambda)$ such that

$$\mathrm{JNR}[D_1(\lambda)] \subseteq \mathrm{JNR}[P(\lambda)] \subseteq \mathrm{JNR}[D_2(\lambda)]. \tag{5}$$

Let $\#\mathrm{NR}[P(\lambda)]$ denote the number of connected components of $\mathrm{NR}[P(\lambda)]$. From the preceding discussion it follows that

$$\#\mathrm{NR}[D_1(\lambda)] \geq \#\mathrm{NR}[P(\lambda)] \geq \#\mathrm{NR}[D_2(\lambda)]. \tag{6}$$

Observe that the orders of $D_1(\lambda)$ and $D_2(\lambda)$ are irrelevant as long as (5) is satisfied. Moreover, (6) gives an approximation of the number of connected components of $\mathrm{NR}[P(\lambda)]$ that can be quite useful, especially when the order n of $P(\lambda)$ is a lot greater than its degree m.

We do know that the joint numerical range of a diagonal matrix polynomial is a convex polyhedron; see [1] and [9]. Conversely, if $K \subset \mathbb{C}^{m+1}$ is a convex polyhedron with n vertices

$$(c_{k,0}, c_{k,1}, \ldots, c_{k,m}), \quad k = 1, 2, \ldots, n,$$

then $K = \mathrm{JNR}[D(\lambda)]$, where

$$D(\lambda) = \mathrm{diag}(c_{1,m}, \ldots, c_{n,m})\lambda^m + \ldots + \mathrm{diag}(c_{1,1}, \ldots, c_{n,1})\lambda + \mathrm{diag}(c_{1,0}, \ldots, c_{n,0}).$$

LEMMA 4.1. *Let $T \subset \mathbb{C}^m$ be a nonempty compact set and W a subspace of $\mathbb{C}^m$ such that $W \cap K \neq \emptyset$ for every convex polyhedron $K \subset \mathbb{C}^m$ that contains T. Then $\mathrm{Co}\{T\} \cap W \neq \emptyset$.*

Proof. Let $\mathcal{K}$ be the family of all convex polyhedra that contain T. Notice that $\mathcal{K}$ is intersection closed. Since T is compact, we have that

$$\mathrm{Co}\{T\} = \bigcap_{K \in \mathcal{K}} K.$$

Define $\mathcal{F} = \{W \cap K : K \in \mathcal{K}\}$. By assumption, all members of $\mathcal{F}$ are compact, nonempty and convex. Moreover, if $W \cap K_j \in \mathcal{F}$ $(j = 1, 2, \ldots, p)$, we have

$$\bigcap_{j=1}^{p} (W \cap K_j) = W \cap (\bigcap_{j=1}^{p} K_j) \in \mathcal{F}$$

since $\bigcap_{j=1}^{p} K_j \in \mathcal{K}$. Hence by Helly's Theorem (see Danzer, Grunbaum, and Klee [3]), $\bigcap_{K \in \mathcal{K}} (W \cap K) \neq \emptyset$. It follows that $\mathrm{Co}\{T\} \cap W = (\bigcap_{K \in \mathcal{K}} K) \cap W \neq \emptyset$. □

The following result is a generalization of [9, Theorem 1.1].

THEOREM 4.2. *Let $P(\lambda) = A_m\lambda^m + \ldots + A_1\lambda + A_0$ be an $n \times n$ matrix polynomial. Then*

$$\bigcup \mathrm{NR}[D_1(\lambda)] = \mathrm{NR}[P(\lambda)] = \bigcap \mathrm{NR}[D_2(\lambda)], \tag{7}$$

where the union [intersection] is taken over all diagonal matrix polynomials $D_1(\lambda)$ $[D_2(\lambda)]$ of degree m for which $\mathrm{JNR}[D_1(\lambda)] \subseteq \mathrm{JNR}[P(\lambda)] \subseteq \mathrm{JNR}[D_2(\lambda)]$.

Proof. For every diagonal matrix polynomial $D_1(\lambda)$ with
$\mathrm{JNR}[D_1(\lambda)] \subseteq \mathrm{JNR}[P(\lambda)]$ it is clear that $\mathrm{NR}[D_1(\lambda)] \subseteq \mathrm{NR}[P(\lambda)]$. Thus

$$\bigcup_{\mathrm{JNR}[D_1(\lambda)] \subseteq \mathrm{JNR}[P(\lambda)]} \mathrm{NR}[D_1(\lambda)] \subseteq \mathrm{NR}[P(\lambda)].$$

In addition, for every $\mathbf{c} = (c_0, c_1, \ldots, c_m) \in \mathrm{JNR}[P(\lambda)]$ we can take $D_1(\lambda) = c_m I\lambda^m + \ldots + c_1 I\lambda + c_0 I$ so that $\mathrm{JNR}[D_1(\lambda)] = \{\mathbf{c}\}$. Hence

$$\bigcup_{\mathrm{JNR}[D_1(\lambda)] \subseteq \mathrm{JNR}[P(\lambda)]} \mathrm{NR}[D_1(\lambda)] \supseteq \mathrm{NR}[P(\lambda)]$$

and the first equation in (7) holds. For the second equation, notice that for every diagonal polynomial matrix $D_2(\lambda)$ of degree m for which $\mathrm{JNR}[D_2(\lambda)] \supseteq \mathrm{JNR}[P(\lambda)]$, we have $\mathrm{NR}[D_2(\lambda)] \supseteq \mathrm{NR}[P(\lambda)]$. Thus

$$\mathrm{NR}[P(\lambda)] \subseteq \bigcap_{\mathrm{JNR}[D_2(\lambda)] \supseteq \mathrm{JNR}[P(\lambda)]} \mathrm{NR}[D_2(\lambda)].$$

Let now

$$\lambda_0 \in \bigcap_{\mathrm{JNR}[D_2(\lambda)] \supseteq \mathrm{JNR}[P(\lambda)]} \mathrm{NR}[D_2(\lambda)].$$

From the discussion in section 2 (see (4)) it follows that $T_1(\lambda_0) \cap \mathrm{JNR}[D_2(\lambda)] \neq \emptyset$ for every diagonal matrix polynomial $D_2(\lambda)$ satisfying $\mathrm{JNR}[D_2(\lambda)] \supseteq \mathrm{JNR}[P(\lambda)]$. Hence from Lemma 4.1,

$$T_1(\lambda_0) \cap \mathrm{Co}\{\mathrm{JNR}[P(\lambda)]\} \neq \emptyset,$$

and from Proposition 2.1, $T_1(\lambda_0) \cap \mathrm{JNR}[P(\lambda)] \neq \emptyset$. So there exists unit $x_0 \in \mathbb{C}^n$ such that

$$(x_0^* A_m x_0)\lambda_0^m + \ldots + (x_0^* A_1 x_0)\lambda_0 + x_0^* A_0 x_0 = 0,$$

which in turn implies

$$\mathrm{NR}[P(\lambda)] \ \supseteq \bigcap_{\mathrm{JNR}[D_2(\lambda)] \supseteq \mathrm{JNR}[P(\lambda)]} \mathrm{NR}[D_2(\lambda)],$$

completing the proof. □

Observation 4.3. In (7) we can replace $\mathrm{JNR}[P(\lambda)]$ by its convex hull.

Observation 4.4. Theorem 4.2 gives a method of approximation of $\mathrm{NR}[P(\lambda)]$ by numerical ranges of diagonal matrix polynomials. With the help of Proposition 2.1, we can develop a second method as follows. Consider two sequences of convex polyhedra $\{K_t^{(i)}\}$ ($i = 1, 2$ and $t = m+2, m+3, \ldots$) so that $K_t^{(i)}$ has t vertices,

$$K_t^{(1)} \subseteq \mathrm{Co}\{\mathrm{JNR}[P(\lambda)]\} \subseteq K_t^{(2)} \quad (t = m+2, m+3, \ldots),$$

and lastly so that both $K_t^{(i)}$ converge to $\mathrm{Co}\{\mathrm{JNR}[P(\lambda)]\}$ as $t \to \infty$. If we let $D_t^{(i)}(\lambda)$ for $i = 1, 2$ be $t \times t$ diagonal matrix polynomials associated with $K_t^{(i)}$, respectively, then for all $t \geq m+2$,

$$\mathrm{NR}[D_t^{(1)}] \subseteq \mathrm{NR}[P(\lambda)] \subseteq \mathrm{NR}[D_t^{(2)}]$$

and both $\mathrm{NR}[D_t^{(i)}]$ converge to $\mathrm{NR}[P(\lambda)]$ as $t \to \infty$. In other words, $\mathrm{NR}[P(\lambda)]$ can be approximated by numerical ranges of diagonal matrix polynomials of the same degree and appropriately large order.

The results in Binding, Farenick, and Li [2] suggest another method of associating numerical ranges of matrix polynomials to numerical ranges of diagonal matrix polynomials via dilations.

Acknowledgments. We acknowledge with thanks an anonymous referee for generously providing us with the construction of the set M_* in section 3. We also thank Doug Farenick and Chi-Kwong Li for useful discussions, and the former for bringing to our attention Helly's theorem and reference [3].

REFERENCES

[1] P. Binding and C. - K. Li. Joint ranges of Hermitian matrices and simultaneous diagonalization. *Linear Algebra and Its Applications*, 151:157–167, 1991.

[2] P. Binding, D. R. Farenick, and C. - K. Li. A dilation and norm in several variable operator theory. *Canadian Journal of Mathematics*, 47:449–461, 1995.

[3] L. Danzer, B. Grunbaum, and V. Klee. Helly's theorem and its relatives. *Proceedings of Symposia in Pure Mathematics*, 7:101–180, 1963.

[4] S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces.* Academic Press, New York, 1978.

[5] G. A. Isaev. Numerical Range of Operator Pencils and Keldysh Multiple Completeness. *Functional Analysis and Its Applications*, 9:27–30, 1975.

[6] C. - K. Li and L. Rodman. Numerical range of matrix polynomials. *SIAM Journal on Matrix Analysis and Applications*, 15:1256–1265, 1994.

[7] Y. Lyubich. Separation of Roots of Matrix and Operator Polynomials. *Integral Equations and Operator Theory*, 29:52–62, 1998.

[8] Y. Lyubich and A. Markus. Connectivity of level sets of quadratic forms and Hausdorff-Toeplitz type theorems. *Positivity*, 1:239–254, 1997.

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

[10] A. M. Ostrowski. *Solutions of equations in Euclidean and Banach spaces.* Academic Press, New York, 1973.

[11] R. Walker. *Algebraic Curves.* Dover, New York, 1950.