On a compression of normal matrix polynomials

Maria Adam[1] and Panayiotis Psarrakos[2]

Department of Mathematics, National Technical University,
Zografou Campus, Athens 15780, Greece

February 21, 2005

Abstract

In this paper, we study a compression of normal matrices and matrix polynomials with respect to a given vector and its orthogonal complement. The numerical range of this compression satisfies special boundary properties, which are investigated in detail. The characteristic polynomial of the compression is also considered.

Keywords: compression; eigenvalue; matrix polynomial; numerical range

AMS Subject Classifications: 15A18; 15A22; 15A60; 47A12; 47A56

1 Introduction and preliminaries

Let $\mathcal{M}_n$ be the algebra of all $n \times n$ complex matrices, and let

$$L(\lambda) = A_m \lambda^m + A_{m-1} \lambda^{m-1} + \cdots + A_1 \lambda + A_0 \tag{1}$$

be a *matrix polynomial*, where $A_j \in \mathcal{M}_n$ $(j = 0, 1, \ldots, m)$ and λ is a complex variable. The matrix polynomial $L(\lambda)$ is called *selfadjoint* if all its coefficients are Hermitian, and it is said to be *normal* if for every $\mu \in \mathbb{C}$, the matrix $L(\mu)$ is normal. The study of matrix polynomials has a long history, especially in the context of their spectral analysis, which leads to the solutions of the corresponding linear systems of differential (or difference) equations [3, 6, 12].

A scalar $\lambda_0 \in \mathbb{C}$ is said to be an *eigenvalue* of the matrix polynomial $L(\lambda)$ in (1) if the system $L(\lambda_0)x = 0$ has a nonzero solution $x_0 \in \mathbb{C}^n$. This solution x_0 is known as an *eigenvector* of $L(\lambda)$ corresponding to λ_0, and the set of all eigenvalues of $L(\lambda)$ is the *spectrum* of $L(\lambda)$, namely, $\sigma(L(\lambda)) =$

---

[1]E-mail: `maria@math.ntua.gr`
[2]E-mail: `ppsarr@math.ntua.gr`

$\{\lambda \in \mathbb{C} : \det L(\lambda) = 0\}$. The polynomial $\det L(\lambda)$ is called the *characteristic polynomial* of $L(\lambda)$ and its degree is no greater than nm. The *numerical range* of $L(\lambda)$ is defined by

$$W(L(\lambda)) = \{\lambda \in \mathbb{C} : x^* L(\lambda)x = 0, \ x \in \mathbb{C}^n, \ x^*x = 1\},$$

and it is always closed and contains $\sigma(L(\lambda))$. If we consider the linear pencil $I\lambda - A$ for some $A \in \mathcal{M}_n$, then $\sigma(I\lambda - A)$ coincides with the spectrum of A, $\sigma(A)$, and $W(I\lambda - A)$ coincides with the classical *numerical range (field of values)* of A, $F(A) = \{x^* A x : x \in \mathbb{C}^n, x^*x = 1\}$, which is always compact and convex [7]. To facilitate the presentation, some properties of $F(A)$, $W(L(\lambda))$ and normal matrix polynomials that are necessary for understanding these notions follow [7, 10, 11, 13, 16].

(A) If $A = B_1 \oplus B_2$ for some $B_1 \in \mathcal{M}_k$ and $B_2 \in \mathcal{M}_{n-k}$ $(1 \le k < n)$, then $F(A) = \text{convex hull}\{F(B_1) \cup F(B_2)\}$.

(B) $F(A)$ contains the convex hull of the eigenvalues of A. If A is normal, then $F(A) = \text{convex hull}\{\sigma(A)\}$ (the converse is not true for $n > 4$).

(C) If μ is a corner of $W(L(\lambda))$ and there is a unit $x_\mu \in \mathbb{C}^n$ such that μ is a simple zero of the scalar polynomial $x_\mu^* L(\lambda) x_\mu$, then $\mu \in \sigma(L(\lambda))$.

(D) If $1 \le k \le n$ and Y is an $n \times k$ complex matrix of rank k, then $W(Y^* L(\lambda)Y) \subseteq W(L(\lambda))$. The equality holds if $k = n$.

(E) The numerical range of $L(\lambda)$ is bounded if and only if $0 \notin F(A_m)$, and in this case, it has no more than m connected components.

(F) If $L(\lambda)$ is normal and there is a $\mu \in \mathbb{C}$ such that the matrix $L(\mu)$ has n distinct eigenvalues, then there exists a unitary $U \in \mathcal{M}_n$ such that $U^* L(\lambda)U$ is diagonal, i.e., the coefficients of $L(\lambda)$ are simultaneously diagonalizable by unitary similarity.

Given two matrices $A \in \mathcal{M}_n$ and $B \in \mathcal{M}_k$ with $1 \le k < n$, B is said to be a *compression* of A if there exists an $n \times k$ matrix P satisfying $P^*P = I_k$ and $B = P^*AP$. Extending this definition, we say that a $k \times k$ matrix polynomial $R(\lambda)$ $(1 \le k < n)$ is a compression of $L(\lambda)$ in (1) if there exists an $n \times k$ matrix P such that $P^*P = I_k$ and $R(\lambda) = P^*L(\lambda)P$ for all $\lambda \in \mathbb{C}$. Then by Property (D), $W(R(\lambda)) \subseteq W(L(\lambda))$.

Recently, compressions of normal matrices have attracted attention and several results have been obtained [1, 4, 5, 15]. In the remainder of this section, some of these results are briefly presented, and in Section 2, their

2

generalizations to normal matrix polynomials are obtained. Moreover, one of the main results in [5] on the characteristic polynomial of the compression is extended and an illustrative example is given. In Section 3, the special case of normal selfadjoint quadratic matrix polynomials is discussed.

Consider now a normal matrix $A \in \mathcal{M}_n$ and let $\lambda_1, \lambda_2, \ldots, \lambda_k \in \sigma(A)$ $(2 \le k \le n)$ such that for every $i = 1, 2, \ldots, k$, the eigenvalue λ_i does not belong to the interior of the convex hull of $\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_k$. Assume that the points $\lambda_1, \lambda_2, \ldots, \lambda_k$, in this particular order, define a closed convex polygon in $\mathbb{C}$, denoted by $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$, where successive λ_i's are allowed to be equal or co-linear, i.e., $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$ may have less than k vertices and can be degenerated to a line segment or a point. Suppose also that $\{x_1, x_2, \ldots, x_k\}$ is an orthonormal system of eigenvectors of A corresponding to $\lambda_1, \lambda_2, \ldots, \lambda_k$, respectively, and let $W = \mathrm{span}\{x_1, x_2, \ldots, x_k\}$. For any vector

$$ v = \sum_{i=1}^{k} c_i \, x_i \in W \ ; \qquad c_i \in \mathbb{C} \setminus \{0\} \ (i = 1, 2, \ldots, k), $$

let $E_W^\perp(v)$ be the orthogonal complement of $\mathrm{span}\{v\}$ with respect to W. Then

$$ y_i \ = \ \frac{\overline{c}_{i+1}}{\sqrt{|c_i|^2 + |c_{i+1}|^2}} \, x_i - \frac{\overline{c}_i}{\sqrt{|c_i|^2 + |c_{i+1}|^2}} \, x_{i+1} \ ; \quad i = 1, 2, \ldots, k \qquad (2) $$

(where $x_{k+1} = x_1$ and $c_{k+1} = c_1$) are unit vectors and belong to $E_W^\perp(v)$, and for every $i = 1, 2, \ldots, k$,

$$ y_i^* A \, y_i \ = \ \frac{|c_{i+1}|^2 \lambda_i + |c_i|^2 \lambda_{i+1}}{|c_i|^2 + |c_{i+1}|^2} \, . $$

Since $x_1, x_2, \ldots, x_k \in \mathbb{C}^n$ are linearly independent and the coefficients $c_1, c_2, \ldots, c_k$ are nonzero, one can easily see that $y_1, y_2, \ldots, y_{k-1}$ are also linearly independent and form a basis of $E_W^\perp(v)$. If $\{w_1, w_2, \ldots, w_{k-1}\}$ is an orthonormal basis of $E_W^\perp(v)$ and we define the $n \times (k-1)$ matrix

$$ P = \begin{bmatrix} w_1 & w_2 & \cdots & w_{k-1} \end{bmatrix}, \qquad (3) $$

then the matrix $B = P^* A P$ is a $(k-1) \times (k-1)$ compression of A and its numerical range is given by

$$ \begin{aligned} F(B) \ &= \ \{(Pz)^* A \, (Pz) : z \in \mathbb{C}^{k-1}, \ z^* z = 1\} \\ &= \ \{x^* A \, x : x \in E_W^\perp(v), \ x^* x = 1\} \\ &\subseteq \ \{x^* A \, x : x \in W, \ x^* x = 1\} \\ &= \ \langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle. \end{aligned} $$

Since $y_1, y_2, \ldots, y_k \in E_W^\perp(v)$, it is clear that denoting $\lambda_{k+1} = \lambda_1$ and $c_{k+1} = c_1$, the points

$$\tau_i = y_i^* A \, y_i = \frac{|c_{i+1}|^2 \lambda_i + |c_i|^2 \lambda_{i+1}}{|c_i|^2 + |c_{i+1}|^2} \;\; ; \quad i = 1, 2, \ldots, k \qquad (4)$$

lie on the boundary of $F(B)$. As a consequence, we have the following known result [1, 4, 5, 15].

Theorem 1 *Suppose $A \in \mathcal{M}_n$ is normal and $\lambda_1, \lambda_2, \ldots, \lambda_k \in \sigma(A)$ such that the convex polygon $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$ is well defined. Let $x_1, x_2, \ldots, x_k$ be the corresponding unit eigenvectors of A, and let P be the $n \times (k-1)$ matrix in (3). Then the numerical range of the $(k-1) \times (k-1)$ compression $B = P^* A P$ lies in $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$, and for every $i = 1, 2, \ldots, k$, $\partial F(B)$ has a common point τ_i with the line segment $\langle \lambda_i, \lambda_{i+1} \rangle$ given by (4).*

It is worth noting that for any $1 \le i < j \le k$, the unit vector

$$y_{i,j} = \frac{\overline{c}_j}{\sqrt{|c_i|^2 + |c_j|^2}} \, x_i - \frac{\overline{c}_i}{\sqrt{|c_i|^2 + |c_j|^2}} \, x_j$$

also belongs to $E_W^\perp(v)$ and the point

$$y_{i,j}^* A \, y_{i,j} = \frac{|c_j|^2 \lambda_i + |c_i|^2 \lambda_j}{|c_i|^2 + |c_j|^2}$$

lies on $F(B) \cap \langle \lambda_i, \lambda_j \rangle$. Furthermore, the methodology discussed above allows the investigation of the case where some λ_i's in $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$ are equal or co-linear. For convenience, we say that a λ_i $(i \in \{1, 2, \ldots, k\})$ is *c-independent* if it is a vertex of the polygon $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$ different from λ_{i-1} and λ_{i+1} (where $\lambda_0 = \lambda_k$ and $\lambda_{k+1} = \lambda_1$). The first statement of the next theorem was proved recently by Gau and Wu [5, Theorem 3], and the other two statements follow readily from the continuity of the numerical range of matrices with respect to the Hausdorff metric.

Theorem 2 *Let $A \in \mathcal{M}_n$ be normal, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be eigenvalues of A such that the polygon $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$ is well defined. Under the hypotheses of Theorem 1, we have the following cases:*

(a) *For every $i \in \{1, 2, \ldots, k\}$ such that λ_i and λ_{i+1} (where $\lambda_{k+1} = \lambda_1$) are c-independent vertices of $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$, the point τ_i in (4) is the only common point of $\partial F(B)$ and $\langle \lambda_i, \lambda_{i+1} \rangle$.*

4

(b) *If there exist $i \in \{1, 2, \ldots, k\}$ and $s \in \{2, 3, \ldots, k-1\}$ such that $\lambda_i, \lambda_{i+1}, \ldots, \lambda_{i+s}$ are co-linear, then $\langle \tau_i, \tau_{i+s-1} \rangle \subseteq \partial F(B) \cap \langle \lambda_i, \lambda_{i+s} \rangle$, where $\lambda_{k+j} = \lambda_j$ and $\tau_{k+j} = \tau_j$ $(j = 1, 2, \ldots)$. If, in addition, λ_i and λ_{i+s} are c-independent, then $\langle \tau_i, \tau_{i+s-1} \rangle = \partial F(B) \cap \langle \lambda_i, \lambda_{i+s} \rangle$.*

(c) *Suppose that there is an index $i \in \{1, 2, \ldots, k\}$ such that λ_i is a vertex of $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$ and appears exactly t times in $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$, and let $\lambda_i = \lambda_{i+1} = \cdots = \lambda_{i+t-1}$. If $t = 1$, then $\partial F(B) \cap \langle \tau_{i-1}, \lambda_i \rangle = \{\tau_{i-1}\}$ (where $\tau_0 = \tau_k$) and $\partial F(B) \cap \langle \lambda_i, \tau_i \rangle = \{\tau_i\}$. If $t > 1$, then $\langle \tau_{i-1}, \lambda_i \rangle \subseteq \partial F(B) \cap \langle \lambda_{i-1}, \lambda_i \rangle$ (where $\lambda_0 = \lambda_k$ and $\tau_0 = \tau_k$) and $\langle \lambda_i, \tau_{i+t-1} \rangle \subseteq \partial F(B) \cap \langle \lambda_i, \lambda_{i+t} \rangle$ (where $\lambda_{k+j} = \lambda_j$ and $\tau_{k+j} = \tau_j$ for $j = 1, 2, \ldots$). In this case, λ_i is a corner of $F(B)$ and an eigenvalue of the compression B.*

Corollary 3 *Any vertex of the convex polygon $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$ that appears more than once in the set $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ is also a corner of $F(B)$ and a common eigenvalue of A and B. Furthermore, if every vertex of $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$ appears more than once in $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$, then $F(B)$ coincides with $\langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$.*

If all the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ are real (this is the case when A is Hermitian), then by Theorems 1 and 2 (b), the numerical range of the compression B coincides with the interval $[\min\{\tau_1, \tau_k\}, \max\{\tau_{k-1}, \tau_k\}]$.

It is worth mentioning that all the results on the numerical range of the described compression (here and in the literature) are essentially based on the existence of mutually orthogonal subspaces of the eigenspaces of A. Suppose that a (general) matrix $A \in \mathcal{M}_n$ is written in the form $A = U^* (D_1 \oplus D_2 \oplus \cdots \oplus D_\xi \oplus T) U$ for some unitary $U \in \mathcal{M}_n$, some diagonal $D_i = \text{diag}\{\lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,k_i}\}$ $(i = 1, 2, \ldots, \xi)$ and an upper triangular $T \in \mathcal{M}_{n-(k_1+\cdots+k_\xi)}$, and assume that the closed convex polygons $\langle \lambda_{1,1}, \lambda_{1,2}, \ldots, \lambda_{1,k_1} \rangle, \langle \lambda_{2,1}, \lambda_{2,2}, \ldots, \lambda_{2,k_2} \rangle, \ldots, \langle \lambda_{\xi,1}, \lambda_{\xi,2}, \ldots, \lambda_{\xi,k_\xi} \rangle$ are well defined and mutually disjoint. Then by applying the above method, keeping in mind $D_1, D_2, \ldots, D_\xi$, we can construct ξ matrices $P_1, P_2, \ldots, P_\xi$ such that for every $1 \leq i, j \leq \xi$ with $i \neq j$, $P_i^* P_i = I_{k_i-1}$, $P_i^* A P_j = 0$ and $F(P_i^* A P_i)$ lies in $\langle \lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,k_i} \rangle$. Moreover, $\partial F(P_i^* A P_i) \cap \partial \langle \lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,k_i} \rangle$ is fully described by Theorem 2. For the $n \times (k_1 + k_2 + \cdots + k_\xi - \xi)$ matrix $P = \begin{bmatrix} P_1 & P_2 & \cdots & P_\xi \end{bmatrix}$, Property (A) implies that

$$
\begin{aligned}
F(P^* A P) &= F\left((P_1^* A P_1) \oplus (P_2^* A P_2) \oplus \cdots \oplus (P_\xi^* A P_\xi) \right) \\
&= \text{convex hull} \left\{ \bigcup_{i=1}^{\xi} F(P_i^* A P_i) \right\}.
\end{aligned}
$$

2 Normal matrix polynomials

Let $L(\lambda) = A_m\lambda^m + \cdots + A_1\lambda + A_0$ be an $n \times n$ matrix polynomial with $\det A_m \neq 0$, and let G be a bounded connected component of its numerical range. Then for any unit vector $x \in \mathbb{C}^n$, the number of zeros of the scalar polynomial $x^*L(\lambda)x$ in G, counting multiplicities, does not depend on x [12, Lemma 26.8], i.e., it is constant. If we denote this constant number by $\mathrm{ind}_L(G)$, then $L(\lambda)$ has exactly $n\,\mathrm{ind}_L(G)$ eigenvalues in G (counting multiplicities) [14, Theorem 2.1]. For an unbounded connected component G of $W(L(\lambda))$, we say that $\mathrm{ind}_L(G) = \kappa$ if for every unit $x \in \mathbb{C}^n$ such that $x^*A_m x \neq 0$, the number of zeros of $x^*L(\lambda)x$ in G is equal to κ.

Suppose now that $L(\lambda)$ is normal, and that for a $\mu \in \mathbb{C}$, the matrix $L(\mu)$ has n distinct eigenvalues. Then by Property (F), there is a unitary matrix $U \in \mathcal{M}_n$ such that

$$U^*L(\lambda)U \;=\; \mathrm{diag}\{d_1(\lambda), d_2(\lambda), \ldots, d_n(\lambda)\} \tag{5}$$

(see [8] for related definitions and results). Since the leading coefficient of $L(\lambda)$ is nonsingular, for every $i = 1, 2, \ldots, n$, the polynomial $d_i(\lambda)$ is of degree m and the ith column of U is a unit eigenvector of $L(\lambda)$ corresponding to the zeros of $d_i(\lambda)$ (in particular, it is a common eigenvector of all the coefficients of $L(\lambda)$). Consequently, if G is a connected component of $W(L(\lambda))$ with $\mathrm{ind}_L(G) = 1$, then $L(\lambda)$ has exactly n eigenvalues in G (one zero of each $d_i(\lambda)$) and the corresponding unit eigenvectors form an orthonormal basis of $\mathbb{C}^n$.

By [16, Proposition 13], the numerical range of a 2×2 diagonal matrix polynomial has no interior, i.e., it is a union of continuous curves. Moreover, by [16, Proposition 14], the boundary of the numerical range of a diagonal matrix polynomial $\mathrm{diag}\{g_1(\lambda), g_2(\lambda), \ldots, g_n(\lambda)\}$ lies on

$$\bigcup_{1 \leq i < j \leq n} W\left(\mathrm{diag}\{g_i(\lambda), g_j(\lambda)\}\right).$$

Theorem 4 *Let $L(\lambda)$ be an $n \times n$ normal matrix polynomial with a nonsingular leading coefficient, satisfying (5), and let G be a connected component of $W(L(\lambda))$ with $\mathrm{ind}_L(G) = 1$. Suppose that for every $i = 1, 2, \ldots, k$ $(2 \leq k \leq n)$, λ_i is the zero of $d_i(\lambda)$ in G (i.e., $\lambda_1, \lambda_2, \ldots, \lambda_k \in \sigma(L(\lambda)) \cap G$) and x_i is the corresponding unit eigenvector of $L(\lambda)$. If P is the $n \times (k-1)$ matrix in (3), then the numerical range of the $(k-1) \times (k-1)$ compression $R(\lambda) = P^*L(\lambda)P$ lies in $W(\mathrm{diag}\{d_1(\lambda), d_2(\lambda), \ldots, d_k(\lambda)\})$, and for every $1 \leq i < j \leq k$ such that $W(\mathrm{diag}\{d_i(\lambda), d_j(\lambda)\}) \cap G$ is bounded, $W(R(\lambda))$*

has a common point with the continuous curve $W(\text{diag}\{d_i(\lambda), d_j(\lambda)\}) \cap G$, that is, the zero of the scalar polynomial $|c_j|^2 d_i(\lambda) + |c_i|^2 d_j(\lambda)$ in G.

Proof Let $1 \le i < j \le k$, and as in the matrix case, consider the unit vector

$$y_{i,j} = \frac{\overline{c}_j}{\sqrt{|c_i|^2 + |c_j|^2}} x_i - \frac{\overline{c}_i}{\sqrt{|c_i|^2 + |c_j|^2}} x_j \in E_W^\perp(v).$$

Then

$$y_{i,j}^* L(\lambda) y_{i,j} = \frac{|c_j|^2 d_i(\lambda) + |c_i|^2 d_j(\lambda)}{|c_i|^2 + |c_j|^2}$$

has exactly one zero in G, which lies on the (bounded) continuous curve

$$W(\text{diag}\{d_i(\lambda), d_j(\lambda)\}) \cap G = \{\lambda \in G : t\, d_i(\lambda) + (1-t) d_j(\lambda), \, 0 \le t \le 1\}$$

(see also the proof of [11, Theorem 2.2]). Hence, the zero of the polynomial $|c_j|^2 d_i(\lambda) + |c_i|^2 d_j(\lambda)$ in G is a common point of the numerical range

$$
\begin{aligned}
W(R(\lambda)) &= \{\lambda \in \mathbb{C} : (Pz)^* L(\lambda)(Pz) = 0, \, z \in \mathbb{C}^{k-1}, \, z^* z = 1\} \\
&= \{\lambda \in \mathbb{C} : x^* L(\lambda) x = 0, \, x \in E_W^\perp(v), \, x^* x = 1\}
\end{aligned}
$$

and the curve $W(\text{diag}\{d_i(\lambda), d_j(\lambda)\}) \cap G$. $\square$

Notice that in the above theorem, if some λ_i's appear more than once in the set $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$, then they also lie in the numerical range of the compression $R(\lambda)$.

For the $(k-1) \times (k-1)$ compression

$$R(\lambda) = B_m \lambda^m + B_{m-1} \lambda^{m-1} + \cdots + B_1 \lambda + B_0,$$

we consider the matrix polynomial

$$
\begin{aligned}
\Pi_R(u, v) &= B_m(u^{m-1} + u^{m-2}v + \cdots + v^{m-1}) + B_{m-1}(u^{m-2} \\
&\quad + u^{m-3}v + \cdots + v^{m-2}) + \cdots + B_2(u+v) + B_1
\end{aligned}
$$

in two complex variables u and v, and next we obtain a sufficient condition for the uniqueness of the intersection points.

Theorem 5 *Suppose that the hypotheses of Theorem 4 hold, and that for some indices $1 \le \zeta_1 < \zeta_2 \le k$, the curve $W(\text{diag}\{d_{\zeta_1}(\lambda), d_{\zeta_2}(\lambda)\}) \cap G$ is bounded, lies on ∂G and its common points with the rest of the curves in the union $\bigcup_{1 \le i < j \le k} (W(\text{diag}\{d_i(\lambda), d_j(\lambda)\}) \cap G)$ are exactly its endpoints.*

Let also τ_{ζ_1,ζ_2} be the zero of $|c_{\zeta_2}|^2 d_{\zeta_1}(\lambda) + |c_{\zeta_1}|^2 d_{\zeta_2}(\lambda)$ in G, and assume that the points $d_{\zeta_1}(\tau_{\zeta_1,\zeta_2})$ and $d_{\zeta_2}(\tau_{\zeta_1,\zeta_2})$ are successive vertices of the convex polygon $F(\mathrm{diag}\{d_1(\tau_{\zeta_1,\zeta_2}), d_2(\tau_{\zeta_1,\zeta_2}), \ldots, d_k(\tau_{\zeta_1,\zeta_2})\})$ that satisfy the conditions of Theorem 2 (a). If for every $\omega \in W(\mathrm{diag}\{d_{\zeta_1}(\lambda), d_{\zeta_2}(\lambda)\}) \cap G$ different from τ_{ζ_1,ζ_2} and for every unit $x \in \mathbb{C}^n$ such that $x^* R(\omega)x = 0$, $x^* \Pi_R(\tau_{\zeta_1,\zeta_2}, \omega)x \neq 0$, then τ_{ζ_1,ζ_2} is the only common point of $\partial W(R(\lambda))$ and the continuous curve $W(\mathrm{diag}\{d_{\zeta_1}(\lambda), d_{\zeta_2}(\lambda)\}) \cap G$.

Proof Without loss of generality, assume that $\zeta_1 = 1$ and $\zeta_2 = 2$, and that the points $d_1(\tau_{1,2})$ and $d_2(\tau_{1,2})$ are successive vertices of $F(\mathrm{diag}\{d_1(\tau_{1,2}), d_2(\tau_{1,2}), \ldots, d_k(\tau_{1,2})\})$ and satisfy the conditions of Theorem 2 (a). Consider a point

$$\omega_1 \in \partial W(R(\lambda)) \cap (W(\mathrm{diag}\{d_1(\lambda), d_2(\lambda)\}) \cap G).$$

Keeping in mind (5), one can see that there exist two unit vectors $y, \hat{y} \in \mathbb{C}^{k-1}$ such that

$$P\,y = U \begin{bmatrix} t \\ e^{\mathrm{i}\theta}\sqrt{1-t^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{and} \quad P\,\hat{y} = U \begin{bmatrix} \hat{t} \\ e^{\mathrm{i}\hat{\theta}}\sqrt{1-\hat{t}^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

for some $t, \hat{t} \in [0,1]$ and $\theta, \hat{\theta} \in [0, 2\pi)$, which satisfy

$$y^* R(\tau_{1,2})y = (P\,y)^* L(\tau_{1,2})(P\,y) = 0$$

and

$$\hat{y}^* R(\omega_1)\hat{y} = (P\,\hat{y})^* L(\omega_1)(P\,\hat{y}) = 0.$$

Hence, the numerical range of the compression $R(\tau_{1,2}) = P^* L(\tau_{1,2})P$ of the matrix $L(\tau_{1,2}) \in \mathcal{M}_n$ contains the origin and the point

$$\hat{y}^* R(\tau_{1,2})\hat{y} = \hat{y}^* \Pi_R(\tau_{1,2}, \omega_1)\hat{y}\,(\tau_{1,2} - \omega_1).$$

Since the conditions of Theorem 2 (a) hold, $\langle d_1(\tau_{1,2}), d_2(\tau_{1,2})\rangle$ is an edge of the numerical range of the matrix $\mathrm{diag}\{d_1(\tau_{1,2}), d_2(\tau_{1,2}), \ldots, d_k(\tau_{1,2})\}$ and contains the origin [13, Theorem 1.1]. By Theorem 2 (a), the origin is the only common point of $\langle d_1(\tau_{1,2}), d_2(\tau_{1,2})\rangle$ and $\partial F(R(\tau_{1,2}))$. Furthermore, the point

$$\begin{aligned} \hat{y}^* R(\tau_{1,2})\hat{y} &= \hat{y}^* \Pi_R(\tau_{1,2}, \omega_1)\hat{y}\,(\tau_{1,2} - \omega_1) \\ &= \begin{bmatrix} \hat{t} & e^{-\mathrm{i}\hat{\theta}}\sqrt{1-\hat{t}^2} \end{bmatrix} \begin{bmatrix} d_1(\tau_{1,2}) & 0 \\ 0 & d_2(\tau_{1,2}) \end{bmatrix} \begin{bmatrix} \hat{t} \\ e^{\mathrm{i}\hat{\theta}}\sqrt{1-\hat{t}^2} \end{bmatrix} \end{aligned}$$

8

also lies on $\langle d_1(\tau_{1,2}), d_2(\tau_{1,2})\rangle$. Thus, it is clear that $\hat{y}^*\Pi_R(\tau_{1,2}, \omega_1)\hat{y} = 0$ or $\omega_1 = \tau_{1,2}$. The proof is complete. $\square$

Next we obtain the characteristic polynomial of the compression $R(\lambda)$, generalizing [5, Theorem 1].

Theorem 6 *Let $L(\lambda)$ be an $n \times n$ normal matrix polynomial, satisfying (5). If $x_1, x_2, \ldots, x_n$ are the corresponding unit eigenvectors of $L(\lambda)$, $v = \sum_{i=1}^{k} c_i x_i$ $(2 \leq k \leq n,\ c_1, c_2, \ldots, c_k \neq 0)$ and P is the $n \times (k-1)$ matrix in (3), then the characteristic polynomial of the $(k-1) \times (k-1)$ compression $R(\lambda) = P^* L(\lambda) P$ is given by*

$$\det R(\lambda) = \sum_{j=1}^{k} \frac{|c_j|^2}{|c_1|^2 + |c_2|^2 + \cdots + |c_k|^2}\, d_1(\lambda)\cdots d_{j-1}(\lambda)\, d_{j+1}(\lambda)\cdots d_k(\lambda).$$

Proof Consider the subspace $W_0 = \left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} \in \mathbb{C}^n : x \in \mathbb{C}^k,\ 0 \in \mathbb{C}^{n-k} \right\}$ of $\mathbb{C}^n$. Then recalling (5), observe that

$$P^* L(\lambda) P = (U^* P)^* \mathrm{diag}\{d_1(\lambda), d_2(\lambda), \ldots, d_n(\lambda)\}\, (U^* P),$$

where the columns of the matrix $U^* P$ form an orthogonal basis of $E_{W_0}^{\perp}(U^* v)$ $\equiv E_{W_0}^{\perp}\left([\, c_1,\ \ldots,\ c_k,\ 0,\ \ldots,\ 0\,]^T\right)$. Moreover, if $L(\lambda)$ is diagonal and $k < n$, then the last $n-k$ rows of P are zero. Hence, without loss of generality, we may assume that $k = n$, and that the matrix polynomial $L(\lambda)$ is diagonal, i.e., $L(\lambda) = \mathrm{diag}\{d_1(\lambda), d_2(\lambda), \ldots, d_n(\lambda)\}$ and $\{x_1 = \mathrm{e}_1,\ x_2 = \mathrm{e}_2,\ \ldots,\ x_n = \mathrm{e}_n\}$ is the standard basis of $\mathbb{C}^n$.

Denote by $P(j)$ the $(n-1) \times (n-1)$ submatrix of P obtained by striking out the jth row $(j = 1, 2, \ldots, n)$. Then by the Binet-Cauchy formula for the determinant of the product of three matrices [2, p. 86],

$$
\begin{aligned}
\det R(\lambda) &= \det\left(P^* \begin{bmatrix} d_1(\lambda) & 0 & \cdots & 0 \\ 0 & d_2(\lambda) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n(\lambda) \end{bmatrix} P \right) \\
&= \sum_{j=1}^{n} |\det P(j)|^2 d_1(\lambda)\cdots d_{j-1}(\lambda)\, d_{j+1}(\lambda)\cdots d_n(\lambda).
\end{aligned}
$$

For the unit vector $\hat{v} = v/\|v\|_2 = \frac{1}{\sqrt{|c_1|^2 + |c_2|^2 + \cdots + |c_n|^2}}\, [\, c_1,\ c_2,\ \ldots,\ c_n\,]^T$, the set $\{w_1, w_2, \ldots, w_{n-1}, \hat{v}\}$ is an orthonormal basis of $\mathbb{C}^n$ and

$$\left| \det\begin{bmatrix} w_1 & w_2 & \cdots & w_{n-1} & \hat{v} \end{bmatrix} \right| = 1.$$

9

For any $j = 1, 2, \ldots, n$, $\{w_1, w_2, \ldots, w_{n-1}, e_j\}$ is a basis of $\mathbb{C}^n$, and thus,

$$
\begin{aligned}
\det \begin{bmatrix} w_1 & w_2 & \cdots & w_{n-1} & \hat{v} \end{bmatrix} &= \det \begin{bmatrix} w_1 & w_2 & \cdots & w_{n-1} & \chi_j\, e_j \end{bmatrix} \\
&= \chi_j \, \det P(j)
\end{aligned}
$$

for some nonzero $\chi_j \in \mathbb{C}$. As a consequence, $|\det P(j)| = |\chi_j|^{-1}$. Moreover, the vectors $y_1, y_2, \ldots, y_{n-1}$ in (2) and e_j also form a basis of $\mathbb{C}^n$, and

$$
\hat{v} = \sum_{i=1}^{n} \frac{c_i}{\sqrt{|c_1|^2 + |c_2|^2 + \cdots + |c_n|^2}} \, e_i = \sum_{i=1}^{n-1} \psi_i\, y_i + \chi_j\, e_j \qquad (6)
$$

for some $\psi_1, \psi_2, \ldots, \psi_{n-1} \in \mathbb{C}$. Straightforward computations imply

$$
\psi_1 = \frac{c_1 \sqrt{|c_1|^2 + |c_2|^2}}{\overline{c}_2 \sqrt{|c_1|^2 + |c_2|^2 + \cdots + |c_n|^2}}
$$

when $j \geq 2$,

$$
\psi_i = \frac{(|c_1|^2 + \cdots + |c_i|^2)\sqrt{|c_i|^2 + |c_{i+1}|^2}}{\overline{c}_i\, \overline{c}_{i+1} \sqrt{|c_1|^2 + |c_2|^2 + \cdots + |c_n|^2}} \quad ; \quad i = 2, 3, \ldots, j-1
$$

when $j \geq 3$,

$$
\psi_{n-1} = -\frac{c_n \sqrt{|c_{n-1}|^2 + |c_n|^2}}{\overline{c}_{n-1} \sqrt{|c_1|^2 + |c_2|^2 + \cdots + |c_n|^2}}
$$

when $j \leq n-1$, and

$$
\psi_i = -\frac{(|c_{i+1}|^2 + \cdots + |c_n|^2)\sqrt{|c_i|^2 + |c_{i+1}|^2}}{\overline{c}_i\, \overline{c}_{i+1} \sqrt{|c_1|^2 + |c_2|^2 + \cdots + |c_n|^2}} \quad ; \quad i = n-2, n-3, \ldots, j
$$

when $j \leq n-2$. Then by substituting to the jth equation of the system (6) (i.e., the one that corresponds to the jth entry of $\hat{v}$), we obtain that

$$
\chi_j = \frac{\sqrt{|c_1|^2 + |c_2|^2 + \cdots + |c_n|^2}}{\overline{c}_j} \, ,
$$

and the proof is complete. $\square$

Let now $A_1\lambda + A_0$ $(\det A_1 \neq 0)$ be an $n \times n$ normal linear pencil with (finite) spectrum

$$
\sigma(A_1\lambda + A_0) = \left\{ \lambda_1 = -\frac{\mu_1}{\nu_1}, \, \lambda_2 = -\frac{\mu_2}{\nu_2}, \, \ldots, \, \lambda_n = -\frac{\mu_n}{\nu_n} \right\},
$$

and suppose that there is a unitary matrix $U \in \mathcal{M}_n$ such that

$$U^*(A_1\lambda + A_0)U \;=\; \mathrm{diag}\{\nu_1\lambda + \mu_1, \nu_2\lambda + \mu_2, \ldots, \nu_n\lambda + \mu_n\}.$$

By [16, Proposition 13 and Corollary 15], the numerical range of a 2×2 diagonal linear pencil $\mathrm{diag}\{\nu\lambda + \mu, \hat{\nu}\lambda + \hat{\mu}\}$ lies either on a straight line or on a circle. Thus, if $W(\mathrm{diag}\{\nu\lambda + \mu, \hat{\nu}\lambda + \hat{\mu}\})$ is bounded (i.e., $0 \notin \langle \nu, \hat{\nu} \rangle$), then it is a line segment or a circular arc. A set of scalars

$$\left\{ \hat{\lambda}_1 = -\frac{\hat{\mu}_1}{\hat{\nu}_1}, \; \hat{\lambda}_2 = -\frac{\hat{\mu}_2}{\hat{\nu}_2}, \; \ldots, \; \hat{\lambda}_k = -\frac{\hat{\mu}_k}{\hat{\nu}_k} \right\}$$

is said to satisfy the *boundary order property* if the boundary of $W(\mathrm{diag}\{\hat{\nu}_1\hat{\lambda} + \hat{\mu}_1, \hat{\nu}_2\hat{\lambda} + \hat{\mu}_2, \ldots, \hat{\nu}_k\hat{\lambda} + \hat{\mu}_k\})$ is bounded and coincides with the union

$$\bigcup_{i=1,2,\ldots,k} W(\mathrm{diag}\{\hat{\nu}_i\hat{\lambda} + \hat{\mu}_i, \hat{\nu}_{i+1}\hat{\lambda} + \hat{\mu}_{i+1}\}),$$

where $\hat{\nu}_{k+1} = \hat{\nu}_1$, $\hat{\mu}_{k+1} = \hat{\mu}_1$, and the curves $W(\mathrm{diag}\{\hat{\nu}_1\hat{\lambda} + \hat{\mu}_1, \hat{\nu}_2\hat{\lambda} + \hat{\mu}_2\})$, $W(\mathrm{diag}\{\hat{\nu}_2\hat{\lambda} + \hat{\mu}_2, \hat{\nu}_3\hat{\lambda} + \hat{\mu}_3\})$, $\ldots$, $W(\mathrm{diag}\{\hat{\nu}_k\hat{\lambda} + \hat{\mu}_k, \hat{\nu}_1\hat{\lambda} + \hat{\mu}_1\})$ are successive parts of $\partial W(\mathrm{diag}\{\hat{\nu}_1\hat{\lambda} + \hat{\mu}_1, \hat{\nu}_2\hat{\lambda} + \hat{\mu}_2, \ldots, \hat{\nu}_k\hat{\lambda} + \hat{\mu}_k\})$. Since $\partial W(\mathrm{diag}\{\hat{\nu}_1\hat{\lambda} + \hat{\mu}_1, \hat{\nu}_2\hat{\lambda} + \hat{\mu}_2, \ldots, \hat{\nu}_k\hat{\lambda} + \hat{\mu}_k\})$ is bounded, the numerical range $W(\mathrm{diag}\{\hat{\nu}_1\hat{\lambda} + \hat{\mu}_1, \hat{\nu}_2\hat{\lambda} + \hat{\mu}_2, \ldots, \hat{\nu}_k\hat{\lambda} + \hat{\mu}_k\})$ is connected. Hence, by Theorems 4 and 5, and observing that for $R(\lambda) = B_1\lambda + B_0$, the matrix polynomial $\Pi_R(u,v)$ coincides with the matrix B_1, we have the following corollaries.

Corollary 7 *Suppose* $\lambda_1 = -\frac{\mu_1}{\nu_1}$, $\lambda_2 = -\frac{\mu_2}{\nu_2}$, $\ldots$, $\lambda_k = -\frac{\mu_k}{\nu_k} \in \sigma(A_1\lambda + A_0)$ *($2 \le k \le n$) satisfy the boundary order property and* $x_1, x_2, \ldots, x_k$ *are the corresponding unit eigenvectors of* $A_1\lambda + A_0$. *For any vector*

$$v = \sum_{i=1}^{k} c_i x_i \;\; ; \qquad c_i \in \mathbb{C} \setminus \{0\} \;\; (i = 1, 2, \ldots, k)$$

with $|c_i|^2\nu_{i+1} + |c_{i+1}|^2\nu_i \neq 0$ *($i = 1, 2, \ldots, k-1$) and* $|c_k|^2\nu_1 + |c_1|^2\nu_k \neq 0$, *consider the* $n \times (k-1)$ *matrix* P *in (3). Then the numerical range of the* $(k-1) \times (k-1)$ *linear pencil* $B_1\lambda + B_0 = P^*(A_1\lambda + A_0)P$ *lies in* $W(\mathrm{diag}\{\nu_1\lambda + \mu_1, \nu_2\lambda + \mu_2, \ldots, \nu_k\lambda + \mu_k\})$, *and for every* $i = 1, 2, \ldots, k$, *the boundary of* $W(B_1\lambda + B_0)$ *has at least a common point with the continuous curve* $W(\mathrm{diag}\{\nu_i\lambda + \mu_i, \nu_{i+1}\lambda + \mu_{i+1}\})$, *that is, the point*

$$\tau_i \;=\; \frac{|c_{i+1}|^2\nu_i\lambda_i + |c_i|^2\nu_{i+1}\lambda_{i+1}}{|c_i|^2\nu_{i+1} + |c_{i+1}|^2\nu_i} \;=\; -\frac{|c_i|^2\mu_{i+1} + |c_{i+1}|^2\mu_i}{|c_i|^2\nu_{i+1} + |c_{i+1}|^2\nu_i}, \qquad (7)$$

where it is assumed that $\nu_{k+1} = \nu_1$, $\mu_{k+1} = \mu_1$, $\lambda_{k+1} = \lambda_1$ *and* $c_{k+1} = c_1$.

Corollary 8 *Suppose that the hypotheses of Corollary 7 hold, and let τ_i $(i = 1, 2, \ldots, k)$ be the intersection points defined by (7). Then for every $i \in \{1, 2, \ldots, k\}$ such that the points $\nu_i \tau_i + \mu_i$ and $\nu_{i+1} \tau_i + \mu_{i+1}$ of $F(\mathrm{diag}\{\nu_1 \tau_i + \mu_1, \nu_2 \tau_i + \mu_2, \ldots, \nu_k \tau_i + \mu_k\})$ satisfy the conditions of Theorem 2 (a), τ_i is the only common point of $\partial W(B_1 \lambda + B_0)$ and $W(\mathrm{diag}\{\nu_i \lambda + \mu_i, \nu_{i+1} \lambda + \mu_{i+1}\})$.*

For the $n \times n$ linear pencil $I\lambda - A$, $\sigma(I\lambda - A) \equiv \sigma(A)$ and $W(I\lambda - A) \equiv F(A)$, and for every $\mu \in \mathbb{C}$, $F(I\mu - A) = \mu - F(A)$. As a consequence, Corollaries 7 and 8 are direct generalizations of Theorems 1 and 2 (a), respectively. Note also that the above corollaries hold even when the numerical range of $A_1\lambda + A_0$ is unbounded, as it is illustrated in the following example.

Example 2 Consider the 3×3 diagonal linear pencil

$$
A_1\lambda + A_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \mathrm{i} & 0 \\ 0 & 0 & -1 - \mathrm{i} \end{bmatrix} \lambda + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
$$

The numerical range $W(A_1\lambda + A_0)$ is unbounded, does not contain the origin and its boundary is sketched in both parts of Figure 1. The eigenvalues of the pencil are marked with o's and satisfy the boundary order property. For the vector $v = [1, \ 1, \ 1]^T$, we construct the 3×2 matrix
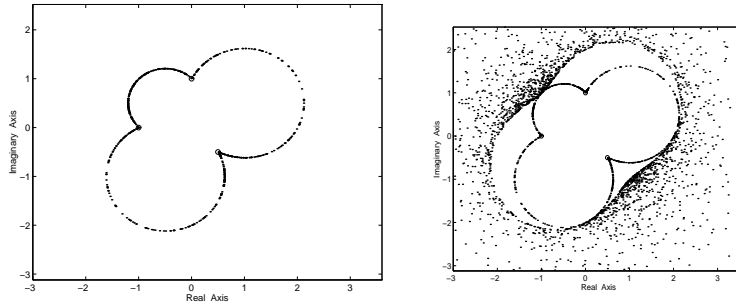

Figure 1: The numerical ranges $W(A_1\lambda + A_0)$ and $W(B_1\lambda + B_0)$.

$$
P \cong \begin{bmatrix} 0.4082 & -0.7071 \\ -0.8165 & 0 \\ 0.4082 & 0.7071 \end{bmatrix},
$$

and observe that $P^*P = I_2$. The numerical range of the linear pencil $B_1\lambda + B_0 = P^*(A_1\lambda + A_0)P$ is drawn in the right part of Figure 1 (in

particular, a few thousand points of it), and $W(B_1\lambda + B_0) \subset W(A_1\lambda + A_0)$. Since
$$W(B_1\lambda + B_0) = \left\{\lambda^{-1} \in \mathbb{C} : \lambda \in F(-B_1),\ \lambda \neq 0\right\}$$
and
$$W(A_1\lambda + A_0) = \left\{\lambda^{-1} \in \mathbb{C} : \lambda \in F(-A_1),\ \lambda \neq 0\right\},$$
one can see that
$$F(B_1\lambda + B_0) \cap W(\mathrm{diag}\{\lambda + 1, \mathrm{i}\,\lambda + 1\}) = -1 + \mathrm{i},$$
$$F(B_1\lambda + B_0) \cap W(\mathrm{diag}\{\lambda + 1, (-1 - \mathrm{i})\lambda + 1\}) = -\mathrm{i}\,2,$$
$$F(B_1\lambda + B_0) \cap W(\mathrm{diag}\{\mathrm{i}\,\lambda + 1, (-1 - \mathrm{i})\lambda + 1\}) = 2,$$
confirming Corollaries 7 and 8. Furthermore, the characteristic polynomial of the compression is $\det(B_1\lambda + B_0) = (-\mathrm{i}/3)\lambda^2 + 1$, verifying Theorem 6.

3   The selfadjoint quadratic case

Selfadjoint quadratic matrix polynomials of the form $S(\lambda) = I\lambda^2 + A_1\lambda + A_0$ arise in many applications [6], and their numerical range is symmetric with respect to the real axis and strongly connected with the numerical range of their coefficients [9].

Theorem 9 [9, Theorem 3] *Let* $S(\lambda) = I\lambda^2 + A_1\lambda + A_0$ *be an* $n \times n$ *selfadjoint matrix polynomial. Then* $\partial W(S(\lambda)) \backslash \mathbb{R}$ *coincides with*

$$\left\{\lambda \in \mathbb{C}\backslash\mathbb{R} : \lambda^2 + a_1\lambda + a_0 = 0,\ a_0 + \mathrm{i}\,a_1 \in \partial F(A_0 + \mathrm{i}\,A_1),\ a_0, a_1 \in \mathbb{R}\right\} =$$

$$\left\{-\frac{a_1}{2} \pm \mathrm{i}\,\frac{\sqrt{4a_0 - a_1^2}}{2} : a_0 + \mathrm{i}\,a_1 \in \partial F(A_0 + \mathrm{i}\,A_1),\ a_0, a_1 \in \mathbb{R},\ 4a_0 > a_1^2\right\}.$$

It is clear that the nonreal part of the boundary of $W(S(\lambda))$ is the union of the images of the part of $\partial F(A_0 + \mathrm{i}\,A_1)$ "inside" the parabola $\mathcal{D} = \{u + \mathrm{i}\,v \in \mathbb{C} : u, v \in \mathbb{R},\ v^2 = 4u\}$ under the mappings

$$a_0 + \mathrm{i}\,a_1 \mapsto -\frac{a_1}{2} + \mathrm{i}\,\frac{\sqrt{4a_0 - a_1^2}}{2} \quad \text{and} \quad a_0 + \mathrm{i}\,a_1 \mapsto -\frac{a_1}{2} - \mathrm{i}\,\frac{\sqrt{4a_0 - a_1^2}}{2}.$$

Suppose now that the matrix polynomial $S(\lambda)$ is normal. This means that the Hermitian matrices A_0 and A_1 have all their eigenspaces in common and there exists a unitary matrix $V \in \mathcal{M}_n$ such that

$$V^*S(\lambda)V = I\lambda^2 + (V^*A_1V)\lambda + V^*A_0V$$

13

$$= I\lambda^2 + \begin{bmatrix} a_{1,1} & 0 & \cdots & 0 \\ 0 & a_{1,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{1,n} \end{bmatrix} \lambda + \begin{bmatrix} a_{0,1} & 0 & \cdots & 0 \\ 0 & a_{0,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{0,n} \end{bmatrix} \tag{8}$$

is a real diagonal matrix polynomial. Then the matrix $A_0 + \mathrm{i}\, A_1$ is normal, $V^*(A_0 + \mathrm{i}\, A_1)V = \mathrm{diag}\{a_{0,1} + \mathrm{i}\, a_{1,1}, a_{0,2} + \mathrm{i}\, a_{1,2}, \ldots, a_{0,n} + \mathrm{i}\, a_{1,n}\}$ and

$$\sigma(S(\lambda)) = \left\{ \lambda \in \mathbb{C} : \lambda^2 + a_{1,i}\lambda + a_{0,i} = 0, \; i = 1, 2, \ldots, n \right\}.$$

By the above discussion and Theorems 1 and 2 (a), we have the following.

Theorem 10 *Suppose* $S(\lambda) = I\lambda^2 + A_1\lambda + A_0$ *is a normal selfadjoint quadratic matrix polynomial, satisfying (8), and*

$$a_{0,1} + \mathrm{i}\, a_{1,1}, \; a_{0,2} + \mathrm{i}\, a_{1,2}, \; \ldots, \; a_{0,k} + \mathrm{i}\, a_{1,k} \in \sigma(A_0 + \mathrm{i}\, A_1) \quad (2 \le k \le n).$$

Let $x_1, x_2, \ldots, x_k$ *be the corresponding unit eigenvectors of* $A_0 + \mathrm{i}\, A_1$, *and let* P *be the* $n \times (k-1)$ *matrix in (3). Then the numerical range of the* $(k-1) \times (k-1)$ *compression* $Q(\lambda) = P^*S(\lambda)P$ *lies in* $W(S(\lambda))$, *and for every* $1 \le i < j \le k$ *such that* $\langle a_{0,i} + \mathrm{i}\, a_{1,i}, a_{0,j} + \mathrm{i}\, a_{1,j} \rangle$ *is an edge of* $F(\mathrm{diag}\{a_{0,1} + \mathrm{i}\, a_{1,1}, a_{0,2} + \mathrm{i}\, a_{1,2}, \ldots, a_{0,k} + \mathrm{i}\, a_{1,k}\})$ *"inside" the parabola* $\mathcal{D}$, $\partial W(Q(\lambda))$ *has a common point with the curve* $\mathcal{C}_{i,j} = \{\lambda \in \mathbb{C} : \lambda^2 + (t\, a_{1,i} + (1-t)a_{1,j})\lambda + t\, a_{0,i} + (1-t)a_{0,j} = 0, \, \mathrm{Im}\lambda > 0, \, t \in [0,1]\} \subset \partial W(S(\lambda)) \setminus \mathbb{R}$ *(resp., with* $\overline{\mathcal{C}}_{i,j}$*), that is,*

$$\xi_{i,j} = \frac{1}{2}\left(-\frac{|c_j|^2 a_{1,i} + |c_i|^2 a_{1,j}}{|c_i|^2 + |c_j|^2} + \mathrm{i}\sqrt{\frac{4(|c_j|^2 a_{0,i} + |c_i|^2 a_{0,j})}{|c_i|^2 + |c_j|^2} - \left(\frac{|c_j|^2 a_{1,i} + |c_i|^2 a_{1,j}}{|c_i|^2 + |c_j|^2}\right)^2} \right)$$

(resp., $\overline{\xi}_{i,j}$*). If, in addition,* $a_{0,i} + \mathrm{i}\, a_{1,i}$ *and* $a_{0,j} + \mathrm{i}\, a_{1,j}$ *are c-independent vertices of* $F(\mathrm{diag}\{a_{0,1} + \mathrm{i}\, a_{1,1}, a_{0,2} + \mathrm{i}\, a_{1,2}, \ldots, a_{0,k} + \mathrm{i}\, a_{1,k}\})$, *then* $\xi_{i,j}$ *(resp.,* $\overline{\xi}_{i,j}$*) is the only common point of* $\partial W(Q(\lambda))$ *and* $\mathcal{C}_{i,j}$ *(resp., and* $\overline{\mathcal{C}}_{i,j}$*).*

Acknowledgments

References

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

[2] A.C. Aitken, *Determinants and Matrices*, Oliver and Boyd, Edinburgh, 1959.

[3] F. Gantmacher, *The Theory of Matrices*, Chelsea, New York, 1959.

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

[5] H.-L. Gau and P.Y. Wu, Numerical range of a normal compression, *Linear and Multilinear Algebra*, to appear.

[6] I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials*, Academic Press, New York, 1982.

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

[8] P. Lancaster and P. Psarrakos, Normal and seminormal eigenvalues of matrix functions, *Integral Equations and Operator Theory*, **41** (2001), pp. 331-342.

[9] P. Lancaster and P. Psarrakos, The numerical range of selfadjoint quadratic matrix polynomials, *SIAM J. Matrix Anal. Appl.*, **23** (2001/02), pp. 615-631.

[10] C.-K. Li, C-numerical ranges and C-numerical radii, *Linear and Multilinear Algebra*, **37** (1994), pp. 51-82.

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

[12] A.S. Markus, *Introduction to the Spectral Theory of Polynomial Operator Pencils*, Amer. Math. Society, Providence, Translations of Math. Monographs, Vol. 71, 1988.

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

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

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

[16] H. Nakazato and P. Psarrakos, On the shape of numerical range of matrix polynomials, *Linear Algebra Appl.*, **338** (2001), pp. 105-123.