Pergamon

Computers Math. Applic. Vol. 31, No. 4/5, pp. 41–47, 1996

0898-1221/96 $15.00 + 0.00

0898-1221(95)00214-6

Geometrical Properties of Numerical Range of Matrix Polynomials

J. MAROULAS AND P. PSARRAKOS
Department of Mathematics, National Technical University
Zografou Campus, Athens 15773, Greece

Abstract—Through the linearization of a matrix polynomial $P(\lambda)$, the symmetry and the sharp points of the numerical range $w(P(\lambda))$ are studied.

Keywords—Numerical range, Matrix polynomials, Linearization, Sharp points.

1. INTRODUCTION

Let M_n be the algebra of all $n \times n$ complex matrices. Suppose that

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

is a matrix polynomial, where $A_l \in M_n$ and λ is a complex variable. The set

$$w(P(\lambda)) = \{\lambda \in \mathbb{C} : x^*P(\lambda)x = 0, \text{ for any } x \in \mathbb{C}^n \text{ such that } x^*x = 1\} \tag{1}$$

is called the numerical range of $P(\lambda)$. If λ_0 is an eigenvalue of $P(\lambda)$, then $\det P(\lambda_0) = 0$, and there exists a nonzero, unitary vector x_0 such that

$$P(\lambda_0)x_0 = [0 \quad 0 \quad \dots \quad 0]^\top, \qquad \text{i.e., } x_0^*P(\lambda_0)x_0 = 0.$$

Hence the eigenvalues of $P(\lambda)$ belong to $w(P(\lambda))$.

For the special case $P(\lambda) = I\lambda - A$, we have

$$w(P(\lambda)) = \{x^*Ax : \text{for } x^*x = 1\},$$

i.e., $w(P(\lambda))$ coincides with the numerical range $w(A)$ of the matrix A. In (1) we consider the restriction $\|x\| = 1$ only for the simplicity of the definition of $w(P(\lambda))$.

It is well known [1] that the $nm \times nm$ matrix

$$L(\lambda) = \lambda \begin{bmatrix} I_n & 0 & \dots & 0 \\ 0 & I_n & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & A_m \end{bmatrix} - \begin{bmatrix} 0 & I & 0\dots & 0 \\ 0 & 0 & I\dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -A_0 & -A_1 & \dots & -A_{m-1} \end{bmatrix} \tag{2}$$

is the linearization of $P(\lambda)$, since

$$E(\lambda)\,L(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} F(\lambda),$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX with

$$F(\lambda) = \begin{bmatrix} I & 0 & 0\dots & 0 \\ -\lambda I & I & 0\dots & 0 \\ \vdots & \ddots & & \\ 0 & \dots & -\lambda I & I \end{bmatrix}, \qquad E(\lambda) = \begin{bmatrix} E_{m-1}(\lambda) & E_{m-2}(\lambda) & E_1(\lambda) & I \\ -I & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & -I & 0 \end{bmatrix},$$

and $E_0(\lambda) = A_m$, $E_{r+1}(\lambda) = \lambda E_r(\lambda) + A_{m-r-1}$ for $r = 0, 1, 2, \dots, m-2$.

Some properties of $w(P(\lambda))$ and the motivation for the study of numerical ranges of matrix polynomials can be found in [2].

The pencil of (2) gives the motivation for a further study of the relationship between the geometrical properties of $w(P(\lambda))$ and the algebraic and analytic properties of $P(\lambda)$.

In this paper, we study the symmetry of $w(P(\lambda))$ and the relation of $w(P(\lambda))$ with the numerical range of its linearization $L(\lambda)$. Furthermore, we give a definition of sharp points of $W(P(\lambda))$ and we present some properties of these points and of the numerical range of Jordan matrices. Finally is given an algorithm for the illustration of $w(A)$, through the MATLAB, which also can be used for the companion matrix of monic matrix polynomials.

2. GENERAL PROPERTIES

In this section we will formulate some properties of the numerical range of matrices and matrix polynomials.

PROPOSITION 2.1. *If the elements a_{k1} of a matrix A belong to the line $\varepsilon : re^{i\theta_0}$, $r \in \mathbb{R}$ for $\theta_0 \in [0, 2\pi]$, then $w(A)$ is symmetric with respect to ε.*

PROOF. The elements of the matrix $e^{-i\theta_0}A$ are real numbers. Then for the numerical range $w(e^{-i\theta_0}A)$, it is verified that

$$w(e^{-i\theta_0}A) = w(e^{-i\theta_0}A^\top) = w(e^{-i\theta_0}A^*) = \overline{w(e^{-i\theta_0}A)},$$

i.e., the $w(e^{-i\theta_0}A)$ is symmetric with respect to the axis $0x$. Therefore, the numerical range $w(A) = e^{i\theta_0}w(e^{-i\theta_0}A)$ is symmetric with respect to the line ε. ∎

COROLLARY. *The numerical range of a real matrix A is symmetric with respect to $\mathbb{R}$.*

PROPOSITION 2.2. *If the coefficients A_i of a $P(\lambda)$ are real matrices, then $w(P(\lambda))$ is symmetric with respect to $\mathbb{R}$.*

PROOF. Let $\lambda_0 \in w(P(\lambda))$, then there exists a nonzero vector $x_0 \in \mathbb{C}^n$ with $\|x_0\| = 1$, such that $x_0^* P(\lambda_0) x_0 = 0$. Considering the conjugate of this equation, we have $x_0^\top P(\bar{\lambda}_0)\bar{x}_0 = 0$. For $y_0 = \bar{x}_0$, it is clear that $y_0^* P(\bar{\lambda}_0) y_0 = 0$, i.e., $\bar{\lambda}_0 \in w(P(\lambda))$. ∎

We write the matrix polynomial in the form $P(\lambda) = P_e(\lambda) + P_{ou}(\lambda)$, where

$$P_e(\lambda) = \dots + A_{2k}\lambda^{2k} + \dots + A_2\lambda^2 + A_0,$$
$$P_o(\lambda) = \dots + A_{2k-1}\lambda^{2k-1} + \dots + A_1\lambda.$$

Then the symmetry of $w(P_e(\lambda))$ and $w(P_o(\lambda))$ with respect to the origin is obvious. Moreover, we have:

PROPOSITION 2.3. *The numerical ranges of the linearizations $L_e(\lambda)$ and $L_o(\lambda)$ of $P_e(\lambda)$ and $P_o(\lambda)$, respectively, are symmetric with respect to the origin.*

PROOF. We will examine the case of $L_e(\lambda)$. If m is even, then for any vector $x = [x_1, \dots, x_m]^\top \in \mathbb{C}^{nm}$, we consider $y = [y_1, \dots, y_m]^\top$ such that

$$y_i = \begin{cases} x_i & \text{for } i \text{ odd}, \\ -x_i & \text{for } i \text{ even}. \end{cases}$$

Thus, we obtain

$$
\begin{aligned}
y^* L_e(\lambda) y &= y^* \left(\lambda \begin{bmatrix} I & 0 & \dots & 0 \\ 0 & I & \dots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & A_m \end{bmatrix} - \begin{bmatrix} 0 & I & & & & \mathbf{0} \\ 0 & 0 & I & & & \vdots \\ & & & & \ddots & I \\ -A_0 & 0 & -A_2 & 0 \dots & -A_{m-2} & 0 \end{bmatrix} \right) y \\
&= \lambda y^* \begin{bmatrix} I & & \dots & 0 \\ & I & & \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & A_m \end{bmatrix} y - \left(y_1^* y_2 + \dots + y_{m-1}^* y_n \right) + y_m^* A_0 y_1 + y_m^* A_2 y_3 \\
&\quad + \dots + y_m^* A_{m-2} y_{m-1} \\
&= \lambda x^* \begin{bmatrix} I & & \dots & 0 \\ & I & & \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & A_m \end{bmatrix} x + x_1^* x_2 + \cdots x_{m-1}^* x_m - \left(x_m^* A_0 x_1 + \dots + x_m^* A_{m-2} x_{m-1} \right) \\
&= -x^* L_e(-\lambda) x.
\end{aligned}
$$

Therefore, for every $\lambda \in w(L_e(\lambda))$, we have that $-\lambda \in w(L_e(\lambda))$ and the proof is complete. ∎

Next we present a relationship between the numerical range of a matrix polynomial $P(\lambda)$ and its linearization $L(\lambda)$.

PROPOSITION 2.4. *If $w(L(\lambda))$ is the numerical range of $L(\lambda)$ in (2), then*

$$w(P(\lambda)) \cup \{0\} \subseteq w(L(\lambda)).$$

PROOF. For any $\lambda \in \mathbb{C}$ and $x \in \mathbb{C}^n$, we consider the vector $y \in \mathbb{C}^{nm}$ defined by the formula:

$$y = \begin{bmatrix} I_n \\ \lambda I_n \\ \vdots \\ \lambda^{m-1} I_n \end{bmatrix} x. \tag{3}$$

Then we can see that

$$y^* L(\lambda) y = x^* \left[I_n \quad \bar{\lambda} I_n \quad \dots \quad \bar{\lambda}^{n-1} I_n \right] \begin{bmatrix} 0 \\ \vdots \\ P(\lambda) x \end{bmatrix} = \bar{\lambda}^{m-1} x^* P(\lambda) x.$$

Thus

$$
\begin{aligned}
w(L(\lambda)) &= \{\lambda : \omega^* L(\lambda) \omega = 0, \text{ for any } \omega\} \\
&\supseteq \{\lambda : y^* L(\lambda) y = 0, \text{ with } y \text{ in (3)}\} \\
&= \{\lambda : \bar{\lambda}^{m-1} x^* P(\lambda) x = 0, \text{ for any } x\} \\
&= \{0\} \cup w(P(\lambda)).
\end{aligned}
$$
∎

3. THE SHARP POINTS OF $W(P(\lambda))$

It is well known that $w(P(\lambda))$ is not always connected, and even if it is connected, it is not always convex.

DEFINITION. *A point* $\lambda_0 \in \partial w(P(\lambda))$[1] *is called a sharp point of* $w(P(\lambda))$ *if for a connected component* $w_s(P(\lambda))$ *of* $w(P(\lambda))$ *there exists a disk* $S(\lambda_0, r)$, $r > 0$, *and the angles* θ_1 *and* θ_2 *with* $0 \le \theta_1 < \theta_2 \le 2\pi$ *such that*

$$\mathrm{Re}(e^{i\theta}\lambda_0) = \max\{\mathrm{Re} z : e^{-i\theta}z \in w_s(P(\lambda)) \cap S(\lambda_0, r)\},$$

for all $\theta \in [\theta_1, \theta_2]$.

Evidently, the isolated points of $w(P(\lambda))$ are sharp points. If $P(\lambda) = I\lambda - A$, then $w(P(\lambda))$ is always a compact convex set, and due to

$$e^{-i\theta}z \in w(P(\lambda)) \Longleftrightarrow z \in w(e^{i\theta}A)),$$

the definition is restricted to that for a matrix A [3, p. 50]. Next we give a property of sharp points for monic matrix polynomials (i.e., $A_m = I$).

PROPOSITION 3.1. *If* λ_0 *is a sharp point of* $L(\lambda)$ *in (2) with* $A_m = I$, *then:*

(i) λ_0 *is an eigenvalue of* $P(\lambda)$, *and*
(ii) λ_0 *is a sharp point of* $w(P(\lambda))$.

PROOF. (i) Let λ_0 be a sharp point of $w(L(\lambda))$; then it is a sharp point of $w(C)$, where C is the companion matrix

$$C = \begin{bmatrix} 0 & I & \dots & 0 \\ \vdots & & I & \\ & & \ddots & \\ \vdots & 0 & & I \\ -A_0 & -A_1 & \dots & -A_{m-1} \end{bmatrix}. \tag{4}$$

Thus, λ_0 is an eigenvalue of C, i.e., λ_0 belongs to the spectrum $\sigma(C) \equiv \sigma(P(\lambda))$.

(ii) By Proposition 2.4 and the definition of sharp points, for any disk $S(\lambda_0, r)$ there exist $\theta_1, \theta_2 \in [0, 2\pi]$ such that

$$\begin{aligned} \mathrm{Re}(e^{i\theta}\lambda_0) &= \max\left\{\mathrm{Re} z : z \in w(e^{i\theta}C)\right\} \\ &= \max\left\{\mathrm{Re} z : e^{-i\theta}z \in w(L(\lambda))\right\} \\ &= \max\left\{\mathrm{Re} z : e^{-i\theta}z \in w(L(\lambda)) \cap S(\lambda_0, r)\right\} \\ &\ge \max\left\{\mathrm{Re} z : e^{-i\theta}z \in w(P(\lambda)) \cap S(\lambda_0, r)\right\}, \end{aligned}$$

for any $\theta \in [\theta_1, \theta_2]$.

Since $\lambda_0 \in \sigma(P(\lambda))$ and $\sigma(P(\lambda)) \subseteq w(P(\lambda))$, then $e^{i\theta}\lambda_0 \in \{z \in \mathbb{C} : e^{-i\theta}z \in w(P(\lambda)) \cap S(\lambda_0, r)\}$ and $\mathrm{Re}(e^{i\theta}\lambda_0) \in \{\mathrm{Re} z : e^{-i\theta}z \in w(P(\lambda)) \cap S(\lambda_0, r)\}$ for any $\theta \in [\theta_1, \theta_2]$. Thus, by the relation

$$\mathrm{Re}(e^{i\theta}\lambda_0) = \max\{\mathrm{Re} z : e^{-i\theta}z \in w(P(\lambda)) \cap S(\lambda_0, r)\},$$

it follows that λ_0 is a sharp point of $w(P(\lambda))$. ∎

The inverse of Proposition 3.1 (ii) does not hold. The next example shows that if λ_0 is a sharp point of $P(\lambda)$, we cannot say that λ_0 is also a sharp point of $L(\lambda)$.

EXAMPLE. Let $P(\lambda) = (\lambda - \beta)^2 I - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, where $\beta \in \mathbb{R}$. Then the spectrum of $P(\lambda)$ is $\sigma(P(\lambda)) = \{\beta - 1, \beta, \beta + 1\}$. For $x = \left[t, e^{i\theta}\sqrt{1 - t^2}\right]^{\mathrm{T}} \in \mathbb{C}^2$; $t \in [0, 1]$, $\theta \in [0, 2\pi]$, we can easily see that $w(P(\lambda))$ is the closed interval $[\beta - 1, \beta + 1]$.

[1]It is not necessary to notice that $\lambda_0 \in \partial w(P(\lambda))$ since $w(P(\lambda))$ is a closed set [2].

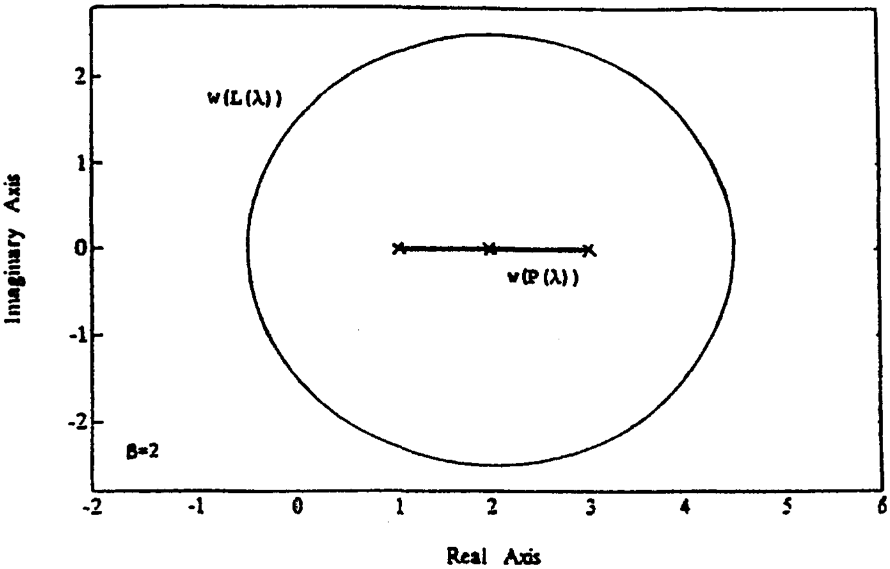

Figure 1.

The elements $\beta - 1$ and $\beta + 1$ are sharp points of $w(P(\lambda))$. Using the algorithm of Section 5, the numerical range of

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1-\beta^2 & 0 & 2\beta & 0 \\ 0 & -\beta^2 & 0 & 2\beta \end{bmatrix}$$

is defined. For $|\beta| \neq 0$, the points $\beta - 1$ and $\beta + 1$ are not sharp points of $w(C)$. Figure 1 illustrates the case for $\beta = 2$.

Finally, using the property that $w(P(\lambda))$ is closed in $\mathbb{C}$ [2] and the notion of the convex hull $\mathrm{Cow}(P(\lambda))$, we state the following proposition.

PROPOSITION 3.2. *For any matrix polynomial $P(\lambda)$, the sharp points of $\mathrm{Cow}(P(\lambda))$ are also sharp points of $w(P(\lambda))$.*

PROOF. Let λ_0 be a sharp point of $\mathrm{Cow}P(\lambda))$. Using Caratheodory's Theorem [4], there exist $r, s, t \in w(P(\lambda))$ such that λ_0 is the center of gravity of the closed triangle $(r, s, t) \subset \mathrm{Cow}(P(\lambda))$. Now, λ_0 coincides with one of the vertices of the triangle, since it is a sharp point. Hence $\lambda_0 \in w(P(\lambda))$, and consequently λ_0 belongs to a connected component $w_s(P(\lambda))$ of $w(P(\lambda))$.

By the definition of sharp points $\exists \theta_1, \theta_2 \in [0, 2\pi]$ such that $\forall \theta \in [\theta_1, \theta_2]$:

$$\begin{aligned} \mathrm{Re}(e^{i\theta}\lambda_0) &= \max\left\{\mathrm{Re}z : e^{-i\theta}z \in \mathrm{Cow}(P(\lambda))\right\} \\ &\geq \max\left\{\mathrm{Re}z : e^{-i\theta}z \in w(P(\lambda))\right\} \\ &\geq \max\left\{\mathrm{Re}z : e^{-i\theta}z \in w_s(P(\lambda)) \cap S(\lambda_0, r)\right\}, \end{aligned}$$

for any disk $S(\lambda_0, r)$. Since $e^{-i\theta}(e^{i\theta}\lambda_0) = \lambda_0 \in w_s(P(\lambda))$, we get that

$$\mathrm{Re}(e^{i\theta}\lambda_0) = \max\left\{\mathrm{Re}z : e^{-i\theta}z \in w_s(P(\lambda)) \cap S(\lambda_0, r)\right\},$$

i.e., λ_0 is a sharp point of $w(P(\lambda))$. ∎

4. THE NUMERICAL RANGE OF JORDAN MATRICES

Consider the $n_k \times n_k$ Jordan block

$$J_k = \lambda_k I + E_k, \qquad \text{where } E_k = \begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & 0 & & \vdots \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

Denoting by ρ_k the spectral radius of $(1/2)(E_k + E_k^{\mathrm{T}})$, we can see that

$$w(J_k) = \{\lambda_k\} + w(E_k) = \{\lambda_k\} + S(0, \rho_k) = S(\lambda_k, \rho_k),$$

where $S(\lambda_k, \rho_k)$ is a disk centered at λ_k with radius ρ_k. Thus for a Jordan matrix

$$J = \operatorname{diag}(J_1, \dots, J_s), \tag{5}$$

we have

$$w(J) = \operatorname{Co}\left\{\bigcup_{k=1}^{s} w(J_k)\right\}. \tag{6}$$

PROPOSITION 4.1. *An eigenvalue λ_i of a Jordan matrix J is a sharp point of $w(J)$ if and only if there are not generalized eigenvectors corresponding to λ_i and* $\lambda_i \notin \operatorname{Co}\left\{\bigcup_{\substack{k=1\\k\neq i}}^{ls} w(J_k)\right\}$.

PROOF. Let λ_i be a sharp point of $w(J)$. Then $\lambda_i \in \partial w(J) \cap \sigma(J)$ which means that λ_i is a normal eigenvalue [3, Theorem 1.6.6]. Thus, does not exist generalized eigenvectors corresponding to λ_i, the Jordan block $J_{\lambda i}$ corresponding to λ_i, has the form $J_{\lambda i} = \lambda_i I$ and $w(J_{\lambda i}) = \{\lambda_i\}$.

If $\lambda_i \in \operatorname{Co}\left\{\bigcup_{\substack{k=1\\k\neq i}}^{ls} w(J_k)\right\}$, then $w(J) = \operatorname{Co}\left(\bigcup_{\substack{k=1\\k\neq i}}^{ls} w(J_k)\right)$. By Proposition 3.2, the λ_i is also a sharp point of $\bigcup_{\substack{k=1\\k\neq i}}^{ls} w(J_k) = \bigcup_{\substack{k=1\\k\neq i}}^{ls} S(\lambda_k, \rho_k)$ which contradicts the property of sharp point.

Conversely, if λ_i has not any generalized eigenvector and does not belong to $\operatorname{Co}\left\{\bigcup_{\substack{k=1\\k\neq i}}^{ls} w(J_k)\right\}$, then the respective Jordan block is $J_{\lambda_i} = \lambda_i I$ and $w(J) = \operatorname{Co}(\{\lambda_i\}) \cup \operatorname{Co}\left(\bigcup_{\substack{k=1\\k\neq i}}^{ls} w(J_k)\right)$.

Therefore, it is clear that $\lambda_i \in \partial w(J)$ and it is a sharp point of $w(J)$. ∎

5. AN ILLUSTRATION OF THE NUMERICAL RANGE

Let $w(A)$ be the numerical range of A. Using Johnson's method [5], we compute r-sided polygons P and Q such that

$$\operatorname{Co}(P) \subseteq w(A) \subseteq \operatorname{Co}(Q),$$

and the boundary of the polygons tends to $\partial w(A)$ as $r \to \infty$. This algorithm can be implemented using MATLAB. The entire process is controlled by a function called NUMRANGE, with input of the matrix A and output of the plots of polygons P and Q. Denoting by p and q the vectors of vertices of polygons P and Q, respectively, we work in the following way:

STEP 1. Choose a partition $\{0 = \theta_0, \theta_1, \theta_2, \dots, \theta_r = 2\pi\}$ of $[0, 2\pi]$.

STEP 2. For $k = 0, 1, \dots, r$, repeat:

(i) Compute eigenvalue $\lambda_k = \lambda_{\max}(H(e^{i\theta_k} A))$.
(ii) Compute unitary eigenvector x_k, associated with λ_k.
(iii) Compute k-vertex of polygon P, $p_k = x_k^* A x_k$.

STEP 3. For $k = 0, 1, \dots, r-1$, compute the vertices of polygon Q:

$$q_k = e^{-i\theta_k}\left[\lambda_k + i\frac{\lambda_k \cos(\theta_{k+1} - \theta_k) - \lambda_{k+1}}{\sin(\theta_{k+1} - \theta_k)}\right].$$

STEP 4. Set $q_r = q_0$.

STEP 5. Compute the difference $EQ = \text{Area}(Q) - \text{Area}(P)$,

$$E = \frac{1}{2}\text{Im}\left[\sum_{k=0}^{r-1}(\bar{q}_k q_{k+1} - \bar{p}_k p_{k+1})\right].$$

STEP 6. Plot polygons P and Q.

We can repeat steps 1–6, taking greater numbers of partition points until the difference E is "small enough." For monic matrix polynomials, we replace A by the companion matrix (4).

REFERENCES

1. I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials*, Chapter 1, Academic Press, London, (1982).
2. C.-K. Li and L. Rodman, Numerical range of matrix polynomials, *SIAM J. of Matrix Analysis* **15**, 1256–1265 (1994).
3. R. Horn and C. Johnson, *Topics in Matrix Analysis*, Chapter 1, Cambridge Univ. Press, New York, (1991).
4. R.E. Edwards, *Functional Analysis: Theory and Applications*, Holt Rinehardt, Winston, (1965).
5. C.R. Johnson, Numerical determination of the field of values of a general complex matrix, *SIAM J. Num. Analysis* **15**, 595–602 (1978).