

ON THE STABILITY RADIUS OF MATRIX POLYNOMIALS

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

Abstract. The stability radius of a matrix polynomial $P(\lambda)$ relative to an open region Ω of the complex plane and its relation to the numerical range of $P(\lambda)$ are investigated. Using an expression of the stability radius in terms of λ on the boundary of Ω and $\|P(\lambda)^{-1}\|_2$, a lower bound is obtained. This bound for the stability radius involves the distances of Ω to the connected components of the numerical range of $P(\lambda)$ and can be applied in conjunction with polygonal approximations of the numerical range. The special case of hyperbolic matrix polynomials is also considered.

Key words : Matrix polynomial, Stability radius, Pseudospectrum, Numerical range

AMS subject classifications: 15A60, 65F15, 47A12, 37C75

Abbreviated title: Stability Radius

1. Introduction. The notions of stability radius and pseudospectrum of a matrix quantify the ‘distance to instability’ relative to a set that contains the spectrum. They are useful tools for investigating the behaviour under perturbations of matrix-based models in dynamical system theory and of algorithms in numerical linear algebra (see e.g., [4, 8, 18, 23]). Furthermore, the pseudospectrum appears in the study of linear operators [2, 22]. The extensions of these notions to matrix polynomials $P(\lambda)$ can play an important role in the solution and perturbation analysis of high order vector differential equations and dynamical systems (see e.g., [7, 12]).

The computational issues that arise regarding the stability radius of matrix polynomials are substantially more complicated than in the case of constant matrices. Methods currently in use (see e.g., [21]) are based either on linearization (resulting in larger order problems), reduction of the order based on projections, or involve the computation of the related functional $\|P(\lambda)^{-1}\|_2$ on a grid.

In this article, we examine the relation of the stability radius to the numerical range of a matrix polynomial. Our approach offers an alternative to the current methods, which can be combined with polygonal approximations of the numerical range for which efficient techniques are available (see Section 4.2). Next section contains basic definitions and notation. In Section 3, we provide expressions of the stability radius relative to an open region Ω in terms of $\|P(\lambda)^{-1}\|_2$, with λ on the boundary of Ω and as some or all of the matrix coefficients are allowed to vary. In Section 4, we obtain a lower bound for the stability radius when the numerical range of $P(\lambda)$ is bounded and lies in Ω . We then consider this bound in the context of polygonal approximations of the numerical range, hyperbolic matrix polynomials and damped vibrating systems.

*Department of Mathematics, National Technical University, Zografou Campus, Athens 15780, Greece (ppsarr@math.ntua.gr).

†Mathematics Department, Washington State University, Pullman, Washington 99164-3113 (tsat@math.wsu.edu). Work partially supported by an NSERC grant.

2. Definitions and notation. Let $\mathcal{M}_n(\mathbb{C})$ ($\mathcal{M}_n(\mathbb{R})$) be the algebra of all $n \times n$ complex (real) matrices and let

$$(2.1) \quad P(\lambda) = A_m \lambda^m + \dots + A_1 \lambda + A_0$$

be a *matrix polynomial*, where $A_j \in \mathcal{M}_n(\mathbb{C})$ ($j = 0, 1, \dots, m$) and $A_m \neq 0$. If all the coefficients of $P(\lambda)$ are Hermitian matrices, then $P(\lambda)$ is called *selfadjoint*. A scalar $\lambda_0 \in \mathbb{C}$ is an *eigenvalue* of $P(\lambda)$ if the equation $P(\lambda_0)x = 0$ has a nonzero solution $x_0 \in \mathbb{C}^n$, known as an *eigenvector* of $P(\lambda)$ corresponding to λ_0 . The set of all eigenvalues of $P(\lambda)$ is known as the *spectrum* of $P(\lambda)$, $\sigma(P) = \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\}$. The *numerical range* of $P(\lambda)$ (see e.g., [11]) is defined as

$$(2.2) \quad W(P) = \{\lambda \in \mathbb{C} : x^* P(\lambda)x = 0 \text{ for some nonzero } x \in \mathbb{C}^n\}.$$

Clearly, $W(P)$ is always closed and contains $\sigma(P)$. When $P(\lambda) = I\lambda - A$, $W(P)$ coincides with the classical *numerical range* (or *field of values*) of the matrix A , $F(A) = \{x^* Ax \in \mathbb{C} : x \in \mathbb{C}^n, x^* x = 1\}$ and $\sigma(P)$ is the spectrum of A , $\sigma(A)$. As shown in [11], $W(P)$ in (2.2) is bounded if and only if $0 \notin F(A_m)$, in which case $W(P)$ has no more than m connected components. One can find more about the geometry of $W(P)$ and the structure of its connected components in [11, 14, 17].

We also need to recall the notions of the *numerical radius* $r(A) = \max\{|\lambda| : \lambda \in F(A)\}$, and the *inner numerical radius* $\tilde{r}(A) = \min\{|\lambda| : \lambda \in F(A)\}$ of a matrix $A \in \mathcal{M}_n(\mathbb{C})$. The open disk centered at μ with radius δ is denoted by $S(\mu, \delta)$ and its closure by $\overline{S(\mu, \delta)}$.

Given an index set $J \subseteq \{0, 1, \dots, m\}$, in this paper, we consider the spectrum of perturbations of the matrix polynomial $P(\lambda)$ in (2.1) of the form

$$(2.3) \quad P_J(\lambda) = (A_m + \Delta_m)\lambda^m + \dots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0,$$

where $\Delta_s = 0$ for all $s \notin J$. With the perturbed polynomial in (2.3) we associate the $n \times n$ matrix polynomial $\Delta_J(\lambda) = \Delta_m \lambda^m + \dots + \Delta_1 \lambda + \Delta_0$ and the $n \times n(m+1)$ complex matrix

$$(2.4) \quad \mathbf{D}_J = [\Delta_m \ \Delta_{m-1} \ \dots \ \Delta_1 \ \Delta_0].$$

Suppose now and for the remainder that Ω is an open region of \mathbb{C} whose boundary, denoted by $\partial\Omega$, is a piecewise smooth curve. The matrix polynomial $P(\lambda)$ is said to be *Ω -stable* if $\sigma(P) \subset \Omega$; in this case, we define the *J -stability radius* of $P(\lambda)$ relative to Ω as

$$\mathcal{R}_J(P, \Omega) = \inf_{\mathbf{D}_J} \{\|\mathbf{D}_J\|_2 : \sigma(P_J) \cap (\mathbb{C} \setminus \Omega) \neq \emptyset\}.$$

That is, $\mathcal{R}_J(P, \Omega)$ is the ‘distance of $P(\lambda)$ to Ω -instability’ when (only) the coefficients of $P(\lambda)$ indexed by J are allowed to vary. The notion of J -stability radius of $P(\lambda)$ is related to the (ϵ, J) -*pseudospectrum* of $P(\lambda)$ for given $\epsilon > 0$, namely,

$$\sigma_{\epsilon, J}(P) = \{\lambda \in \mathbb{C} : \lambda \in \sigma(P_J) \text{ for some } \Delta_J(\lambda) \text{ with } \|\mathbf{D}_J\|_2 \leq \epsilon\},$$

where $P_J(\lambda)$ and \mathbf{D}_J are as in (2.3) and (2.4), respectively. It follows from the definitions that $\sigma_{\epsilon, J}(P) \subset \Omega$ if and only if $\mathcal{R}_J(P, \Omega) > \epsilon$.

Notice that when $J = \{0, 1, \dots, m\}$, the J -stability radius and the (ϵ, J) -pseudospectrum coincide with the notions of stability radius and pseudospectrum of matrix polynomials as found in [16, 5, 21]. These notions, in turn, generalize the corresponding notions for matrices; recall that the *stability radius* of a matrix $A \in \mathcal{M}_n(\mathbb{C})$ relative to Ω is $\mathcal{R}(A, \Omega) = \inf\{\|\Delta\|_2 : \Delta \in \mathcal{M}_n(\mathbb{C}), \sigma(A + \Delta) \cap (\mathbb{C} \setminus \Omega) \neq \emptyset\}$.

3. Stability radius and pseudospectrum. Consider a matrix polynomial $P(\lambda)$ as in (2.1) and its perturbation $P_J(\lambda)$ in (2.3), and let Ω be an open region of \mathbb{C} such that $\sigma(P) \subset \Omega$. Observe that for every $\lambda \in \partial\Omega$, the matrix $P(\lambda)$ is invertible, and that $\det P_J(\lambda) = 0$ if and only if $\det(I + \Delta_J(\lambda)P(\lambda)^{-1}) = 0$. By the definition of J -stability radius and by the continuity of the eigenvalues of $P(\lambda)$ with respect to the entries of the coefficient matrices, it follows that

$$\begin{aligned}
\mathcal{R}_J(P, \Omega) &= \inf_{\lambda \notin \Omega} \left\{ \inf_{\mathbf{D}_J} \{\|\mathbf{D}_J\|_2 : \det P_J(\lambda) = 0\} \right\} \\
&= \inf_{\lambda \in \partial\Omega} \left\{ \inf_{\mathbf{D}_J} \{\|\mathbf{D}_J\|_2 : \det P_J(\lambda) = 0\} \right\} \\
(3.1) \quad &= \inf_{\lambda \in \partial\Omega} \left\{ \inf_{\mathbf{D}_J} \{\|\mathbf{D}_J\|_2 : \det(I + \Delta_J(\lambda)P(\lambda)^{-1}) = 0\} \right\},
\end{aligned}$$

where \mathbf{D}_J is defined in (2.4). Notice that $P(\lambda)^{-1}$ assumes the role of the *resolvent* $(I\lambda - A)^{-1}$ of the matrix A in [4, 8, 18, 22, 23].

Next, we present a basic expression of the J -stability radius of $P(\lambda)$. The proof employs the methodology in [5] (where in our notation, $J = \{0, 1, \dots, m\}$) and contains a construction of the perturbations that attain the corresponding infimum. See also [16, 21] for proofs of related results based on the notion of companion linearization of $P(\lambda)$.

THEOREM 3.1. *Let $P(\lambda) = A_m\lambda^m + \dots + A_1\lambda + A_0$ be an $n \times n$ matrix polynomial with $\det A_m \neq 0$, and let $J \subseteq \{0, 1, \dots, m\}$. If Ω is an open region of \mathbb{C} such that $\sigma(P) \subset \Omega$, then*

$$\mathcal{R}_J(P, \Omega) = \inf \left\{ \frac{1}{\sqrt{\sum_{k \in J} |\lambda|^{2k} \|P(\lambda)^{-1}\|_2}} : \lambda \in \partial\Omega \right\}.$$

Proof. Since $\det A_m \neq 0$, $P(\lambda)$ has nm finite eigenvalues (counting their multiplicities) [7]. Consider a (fixed) $\lambda \in \mathbb{C} \setminus \sigma(P)$ and a matrix polynomial

$$\Delta_J(\lambda) = \Delta_m\lambda^m + \dots + \Delta_1\lambda + \Delta_0 = \mathbf{D}_J [I\lambda^m \dots I\lambda \ I]^T,$$

where \mathbf{D}_J is as in (2.4). Suppose that $\det(I + \Delta_J(\lambda)P(\lambda)^{-1}) = 0$. Then -1 is an eigenvalue of the rational matrix function $\Delta_J(\lambda)P(\lambda)^{-1}$ and thus,

$$1 \leq \|\Delta_J(\lambda)P(\lambda)^{-1}\|_2 \leq \|\Delta_J(\lambda)\|_2 \|P(\lambda)^{-1}\|_2.$$

As a consequence, $\|\Delta_J(\lambda)\|_2 \geq \|P(\lambda)^{-1}\|_2^{-1}$, which implies

$$(3.2) \quad \|\mathbf{D}_J\|_2 \geq \frac{1}{\sqrt{\sum_{k \in J} |\lambda|^{2k} \|P(\lambda)^{-1}\|_2}}.$$

Furthermore, one can construct matrices Δ_s ($s = 0, 1, \dots, m$) for which \mathbf{D}_J attains the above lower bound and $\det(I + \Delta_J(\lambda)P(\lambda)^{-1}) = 0$, as follows. Consider two vectors $x, y \in \mathbb{C}^n$ such that $\|x\|_2 = 1$, $\|P(\lambda)^{-1}x\|_2 = \|P(\lambda)^{-1}\|_2$ and

$$y_j = \frac{\omega_j}{\|P(\lambda)^{-1}\|_2^2}, \quad j = 1, 2, \dots, n,$$

where $\omega = [\omega_1 \ \omega_2 \ \dots \ \omega_n]^T := P(\lambda)^{-1}x$. Define the matrix $Q_0 = -xy^*$ and let either

$$\Delta_s = \frac{\bar{\lambda}^s}{\sum_{k \in J} |\lambda|^{2k}} Q_0 \quad (s \in J) \quad \text{or} \quad \Delta_s = 0 \quad (s \notin J).$$

Then $(I + \Delta_J(\lambda)P(\lambda)^{-1})x = x + Q_0 P(\lambda)^{-1}x = x + Q_0 \omega$ and, as $y^* \omega = 1$,

$$(I + \Delta_J(\lambda)P(\lambda)^{-1})x = x - x y^* \omega = 0.$$

Thus, $\det(I + \Delta_J(\lambda)P(\lambda)^{-1}) = 0$. We also have

$$\begin{aligned} \|\mathbf{D}_J\|_2 &= \sup \left\{ \frac{\|Q_0 (\sum_{k \in J} \bar{\lambda}^k v_k) (\sum_{k \in J} |\lambda|^{2k})^{-1}\|_2}{\sqrt{\sum_{k \in J} \|v_k\|_2^2}} : v_k \in \mathbb{C}^n \setminus \{0\} \right\} \\ &\leq \frac{1}{\sum_{k \in J} |\lambda|^{2k}} \sup_{v_k \neq 0} \left\{ \frac{\|x\|_2 \|y\|_2 \|\sum_{k \in J} \bar{\lambda}^k v_k\|_2}{\sqrt{\sum_{k \in J} \|v_k\|_2^2}} \right\}. \end{aligned}$$

Moreover, by construction,

$$\|y\|_2 = \frac{\|\omega\|_2}{\|P(\lambda)^{-1}\|_2^2} = \frac{\|P(\lambda)^{-1}x\|_2}{\|P(\lambda)^{-1}\|_2^2} = \frac{1}{\|P(\lambda)^{-1}\|_2}$$

and

$$\left\| \sum_{k \in J} \bar{\lambda}^k v_k \right\|_2 \leq \left\| [I \ I\lambda \ \dots \ I\lambda^m] \right\|_2 \left\| [v_0^T \ v_1^T \ \dots \ v_m^T]^T \right\|_2,$$

where $v_k = 0$ for $k \notin J$. Therefore, as $\|x\|_2 = 1$, we have

$$\|\mathbf{D}_J\|_2 \leq \frac{1}{\sqrt{\sum_{k \in J} |\lambda|^{2k}} \|P(\lambda)^{-1}\|_2}.$$

That is, for this special \mathbf{D}_J (which depends on λ), equality holds in (3.2). The proof is complete in view of (3.1) and since $\partial\Omega \cap \sigma(P) = \emptyset$. \square

The proof of the above theorem also yields the following.

THEOREM 3.2. *Let $P(\lambda) = A_m \lambda^m + \dots + A_1 \lambda + A_0$ be an $n \times n$ matrix polynomial with $\det A_m \neq 0$, and let $J \subseteq \{0, 1, \dots, m\}$ and $\epsilon > 0$. Then*

$$\sigma_{\epsilon, J}(P) \setminus \sigma(P) = \left\{ \lambda \in \mathbb{C} \setminus \sigma(P) : \frac{1}{\sqrt{\sum_{k \in J} |\lambda|^{2k}} \|P(\lambda)^{-1}\|_2} \leq \epsilon \right\}.$$

Proof. Consider a (fixed) $\lambda \in \mathbb{C} \setminus \sigma(P)$. If $\lambda \in \sigma_{\epsilon, J}(P)$, then there is an $n \times n$ matrix polynomial $\Delta_J(\lambda) = \Delta_m \lambda^m + \dots + \Delta_1 \lambda + \Delta_0$ such that $\Delta_s = 0$ for $s \notin J$, $\|[\Delta_m \dots \Delta_1 \Delta_0]\|_2 \leq \epsilon$ and $\det(P(\lambda) + \Delta_J(\lambda)) = 0$. Thus, by (3.2),

$$\frac{1}{\sqrt{\sum_{k \in J} |\lambda|^{2k} \|P(\lambda)^{-1}\|_2}} \leq \|[\Delta_m \dots \Delta_1 \Delta_0]\|_2 \leq \epsilon.$$

Conversely, suppose that

$$\frac{1}{\sqrt{\sum_{k \in J} |\lambda|^{2k} \|P(\lambda)^{-1}\|_2}} \leq \epsilon.$$

Then, as in the proof of Theorem 3.1, one can construct a matrix polynomial $\Delta_J(\lambda) = \Delta_m \lambda^m + \dots + \Delta_1 \lambda + \Delta_0$ such that $\Delta_s = 0$ for $s \notin J$, $\|[\Delta_m \dots \Delta_1 \Delta_0]\|_2 \leq \epsilon$ and $\det(P(\lambda) + \Delta_J(\lambda)) = 0$. Thus, $\lambda \in \sigma_{\epsilon, J}(P)$. \square

Due to the continuity of $\sqrt{\sum_{k \in J} |\lambda|^{2k} \|P(\lambda)^{-1}\|_2}$ as a function of λ , the boundary of the (ϵ, J) -pseudospectrum can be described as follows.

COROLLARY 3.3. *Let $P(\lambda) = A_m \lambda^m + \dots + A_1 \lambda + A_0$ be an $n \times n$ matrix polynomial with $\det A_m \neq 0$ and let $\epsilon > 0$. Then*

$$\partial(\sigma_{\epsilon, J}(P) \setminus \sigma(P)) = \left\{ \lambda \in \mathbb{C} \setminus \sigma(P) : \frac{1}{\sqrt{\sum_{k \in J} |\lambda|^{2k} \|P(\lambda)^{-1}\|_2}} = \epsilon \right\}.$$

REMARK 3.4. We have chosen to present our results in terms of the 2-norm due to its association with the notion of singular values (see next section and [3, 9, 19]). However, Theorems 3.1 and 3.2 and Corollary 3.3 (with minor modifications in their proofs) are also valid for p -norms: The quantity

$$\sqrt{\sum_{k \in J} |\lambda|^{2k}}$$

should be replaced by

$$\begin{cases} \sqrt{\sum_{k \in J} |\lambda|^{kp}} & \text{if } p < +\infty \\ \max\{|\lambda|^k : k \in J\} & \text{if } p = +\infty. \end{cases}$$

EXAMPLE 3.5. Consider the matrix polynomial

$$P(\lambda) = I\lambda^2 - \begin{pmatrix} 0 & 0.5 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.25 \end{pmatrix}$$

and the open unit disk $\Omega = S(0, 1)$. The spectrum of $P(\lambda)$ is $\sigma(P) = \{-0.5, 0, 0.5\}$. Let us examine the (ϵ, J) -pseudospectrum of $P(\lambda)$ for $J = \{0, 1, 2\}$. For every $\lambda \in \partial\Omega$, $\sqrt{1 + |\lambda|^2 + |\lambda|^4} = \sqrt{3}$ and

$$P(\lambda)^{-1} = \begin{pmatrix} \lambda^{-2} & 0.5\lambda^{-4} & 0 \\ 0 & \lambda^{-2} & 0 \\ 0 & 0 & (\lambda^2 - 0.25)^{-1} \end{pmatrix}$$

with $1.2808 \leq \|P(\lambda)^{-1}\|_2 \leq 1.3333$. Hence, $\mathcal{R}_J(P, \Omega) = (1.3333\sqrt{3})^{-1} = 0.4330$. That is, $P_J(\lambda) = P(\lambda) + \Delta_2\lambda^2 + \Delta_1\lambda + \Delta_0$ with $\|[\Delta_2 \ \Delta_1 \ \Delta_0]\|_2 < 0.4330$, is Ω -stable. Equivalently, for every $\epsilon < 0.4330$, $\sigma_{\epsilon, J}(P)$ remains in Ω . This is confirmed visually in Figure 3.1, where $\partial\Omega$, $\sigma(P)$ and $\partial\sigma_{\epsilon, J}(P)$ for $\epsilon = 0.04, 0.09, 0.2, 0.3, 0.4$ are sketched (using Corollary 3.3).

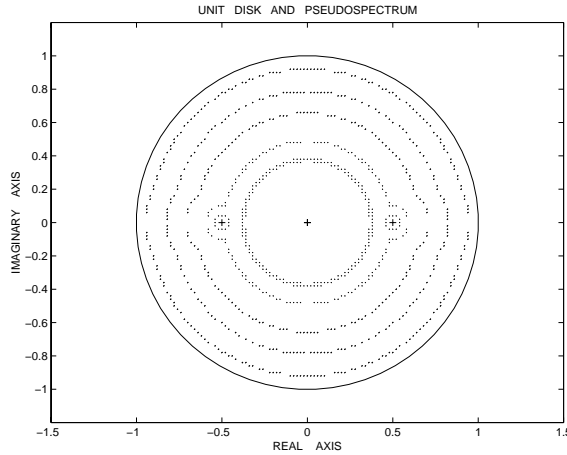


FIG. 3.1. The (ϵ, J) -pseudospectrum of a matrix polynomial.

4. Bounds for the stability radius and applications.

4.1. Stability radius and the connected components. Let $P(\lambda)$ be an $n \times n$ matrix polynomial with bounded numerical range $W(P)$ and let \mathcal{G} be a connected component of $W(P)$. By [12, Lemma 26.8], for every nonzero vector $x \in \mathbb{C}^n$, the number of roots of the equation $x^*P(\lambda)x = 0$ in \mathcal{G} is constant, i.e., it does not depend on x . We denote this constant number by $c(\mathcal{G})$.

First, we obtain a lower bound for $\|P(\lambda)^{-1}\|_2$, $\lambda \in \mathbb{C} \setminus W(P)$ (cf. [6, Lemma 3.2] for operators and [13, Theorem 1] for analytic operator functions).

THEOREM 4.1. *Let $P(\lambda) = A_m\lambda^m + \dots + A_1\lambda + A_0$ be an $n \times n$ matrix polynomial with bounded numerical range $W(P)$ that consists of s connected components $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_s$. Then for every $\lambda \in \mathbb{C} \setminus W(P)$,*

$$\|P(\lambda)^{-1}\|_2 \leq \frac{1}{\tilde{r}(A_m) \prod_{j=1}^s \text{dist}[\lambda, \mathcal{G}_j]^{c(\mathcal{G}_j)}}.$$

Proof. For any nonzero vector $x \in \mathbb{C}^n$, let $\lambda_1(x), \lambda_2(x), \dots, \lambda_m(x)$ be the roots of the equation $x^*P(\lambda)x = 0$. Then for every $\lambda \in \mathbb{C}$,

$$(4.1) \quad \|P(\lambda)x\|_2 \|x\|_2 \geq |x^*P(\lambda)x| = |(x^*A_mx)(\lambda - \lambda_m(x)) \dots (\lambda - \lambda_1(x))|.$$

By the assumption that $W(P)$ is bounded, it follows that $0 \notin F(A_m)$ and thus $\tilde{r}(A_m) > 0$. Since for every nonzero $x \in \mathbb{C}^n$, $\tilde{r}(A_m) \leq |x^*A_mx|/(x^*x)$, (4.1) implies that

$$(4.2) \quad \|P(\lambda)x\|_2 \geq \tilde{r}(A_m) \prod_{j=1}^s \text{dist}[\lambda, \mathcal{G}_j]^{c(\mathcal{G}_j)} \|x\|_2, \quad \lambda \in \mathbb{C}.$$

Since $\det P(\lambda) \neq 0$ for all $\lambda \in \mathbb{C} \setminus \sigma(P)$, we can substitute $x = P(\lambda)^{-1}\tilde{x}$ in (4.2) to obtain

$$\|\tilde{x}\|_2 \geq \tilde{r}(A_m) \prod_{j=1}^s \text{dist}[\lambda, \mathcal{G}_j]^{c(\mathcal{G}_j)} \|P(\lambda)^{-1}\tilde{x}\|_2, \quad \lambda \in \mathbb{C} \setminus W(P)$$

or equivalently,

$$\frac{\|P(\lambda)^{-1}\tilde{x}\|_2}{\|\tilde{x}\|_2} \leq \frac{1}{\tilde{r}(A_m) \prod_{j=1}^s \text{dist}[\lambda, \mathcal{G}_j]^{c(\mathcal{G}_j)}}, \quad \lambda \in \mathbb{C} \setminus W(P).$$

Since the above inequality holds for all nonzero \tilde{x} , the proof is complete. \square

THEOREM 4.2. *Let $P(\lambda) = A_m\lambda^m + \dots + A_1\lambda + A_0$ be an $n \times n$ matrix polynomial such that $W(P)$ is bounded and consists of s connected components $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_s$. If Ω is an open region of \mathbb{C} such that $W(P) \subset \Omega$ and $J \subseteq \{0, 1, \dots, m\}$, then*

$$\mathcal{R}_J(P, \Omega) \geq \inf \left\{ \frac{\tilde{r}(A_m) \prod_{j=1}^s \text{dist}[\lambda, \mathcal{G}_j]^{c(\mathcal{G}_j)}}{\sqrt{\sum_{k \in J} |\lambda|^{2k}}} : \lambda \in \partial\Omega \right\}.$$

Proof. The result follows from Theorems 3.1 and 4.1. \square

4.2. The bound via polygonal approximations. Theorem 4.2 in conjunction with the results in [17] provide the possibility to compute a bound for the J -stability radius of a matrix polynomial when only rough approximations of the numerical ranges of the coefficients are known. We describe and illustrate this technique next.

Let $P(\lambda) = A_m\lambda^m + \dots + A_1\lambda + A_0$ be an $n \times n$ matrix polynomial. Suppose that for each $j = 0, 1, \dots, m$,

$$F(A_j) \subseteq \Psi_j := \text{convex hull} \{ \psi_{j,1}, \psi_{j,2}, \dots, \psi_{j,\xi_j} \}.$$

Note that one can easily construct convex polygons containing the numerical range of a matrix; see [10, pp. 33-37]. Let $\zeta := \prod_{j=0}^m \xi_j$ and consider the $\zeta \times \zeta$ diagonal matrix polynomial

$$D(\lambda) = D_m\lambda^m + \dots + D_1\lambda + D_0,$$

whose coefficients are defined as follows: For each $j = 0, 1, \dots, m$, the matrix

$$D_j := \text{diag}\{d_{j,1}, d_{j,2}, \dots, d_{j,\zeta}\}$$

has $\prod_{k \neq j} \xi_k$ diagonal entries equal to $\psi_{j,\tau}$ for $\tau = 1, 2, \dots, \xi_j$, arranged so that for every selection of points $\{\psi_{0,k_0}, \psi_{1,k_1}, \dots, \psi_{m,k_m}\}$, there exists $\nu \in \{1, 2, \dots, \zeta\}$ such that

$$(d_{0,\nu}, d_{1,\nu}, \dots, d_{m,\nu}) = (\psi_{0,k_0}, \psi_{1,k_1}, \dots, \psi_{m,k_m}).$$

By [17, Theorem 2.1], one deduces that $W(P) \subseteq W(D)$. Consequently, if Ω is an open region that contains $W(D)$ and if $\mu \in \partial\Omega$, then the distance of μ from each connected component of $W(D)$ is not greater than the distance of μ from the corresponding component of $W(P)$. Thus, the bound of Theorem 4.2 is valid when the distances are replaced by distances of λ from connected components of $W(D)$.

The advantages of the above proposed method are the following:

1. It is independent of n .
2. It provides a bound for $\mathcal{R}_J(P, \Omega)$ even when the information given pertains only to the numerical ranges of the coefficients of $P(\lambda)$.
3. The boundary of the numerical range of a diagonal matrix polynomial can be readily computed [15].

We illustrate the above by an example.

EXAMPLE 4.3. Let $\Omega = S(0, 1)$ and consider a matrix polynomial $P(\lambda) = I\lambda^3 + A_2\lambda^2 + A_1\lambda + A_0$ for which the only information given is that $W(P) \subset \Omega$, the coefficients are Hermitian, and that the minimum and maximum eigenvalues of the matrix coefficients satisfy

$$0 \leq \lambda_{\min}(A_2) \leq \lambda_{\max}(A_2) \leq 1/3,$$

$$-1/9 \leq \lambda_{\min}(A_1) \leq \lambda_{\max}(A_1) \leq 1/9,$$

$$-1/27 \leq \lambda_{\min}(A_0) \leq \lambda_{\max}(A_0) \leq 1/9.$$

In the notation used above, $m = 3$, $j \in \{0, 1, 2, 3\}$, $\zeta = 1 \cdot 2^3 = 8$, and $A_3 = I$. According to the method described above, let us consider the (trivial) polygons

$$\Psi_3 = \{1\}, \Psi_2 = [0, 1/3], \Psi_1 = [-1/9, 1/9], \Psi_0 = [-1/27, 1/9]$$

and the 8×8 matrix polynomial $D(\lambda) = D_3\lambda^3 + D_2\lambda^2 + D_1\lambda + D_0$, where $D_3 = I$ and

$$D_2 = \text{diag}\{0, 0, 0, 0, 1/3, 1/3, 1/3, 1/3\},$$

$$D_1 = \text{diag}\{-1/9, -1/9, 1/9, 1/9, -1/9, -1/9, 1/9, 1/9\},$$

$$D_0 = \text{diag}\{-1/27, 1/9, -1/27, 1/9, -1/27, 1/9, -1/27, 1/9\}.$$

Thus, all 8 possible choices $(\psi_{0,k_0}, \dots, \psi_{3,k_3})$ can be formed from entries of the D_j 's as specified by our method. The numerical range of $D(\lambda)$ is the symmetric region outlined

inside the unit disk in Figure 4.1. The curves drawn belong to $W(D)$ and their union contains $\partial W(D)$. From this figure one estimates that $\text{dist}(\partial\Omega, W(D)) = 0.2899$, attained by the left-most real eigenvalue of $D(\lambda)$ (which equals -0.7101). Suppose now, for the sake of illustration, that only the coefficient matrices A_0, A_1, A_2 are perturbed, but not $A_3 = I$; that is $J = \{0, 1, 2\}$. Then, by the discussion preceding this example,

$$\mathcal{R}_J(P, \Omega) \geq \text{dist}(\partial\Omega, W(P))^3 / \sqrt{3} \geq \text{dist}(\partial\Omega, W(D))^3 / \sqrt{3} = 0.2899^3 / \sqrt{3} = 0.0141.$$

Note that this bound is valid for every matrix polynomial $Q(\lambda)$ of degree 3 such that $W(Q) \subseteq W(D) \subseteq \Omega$.

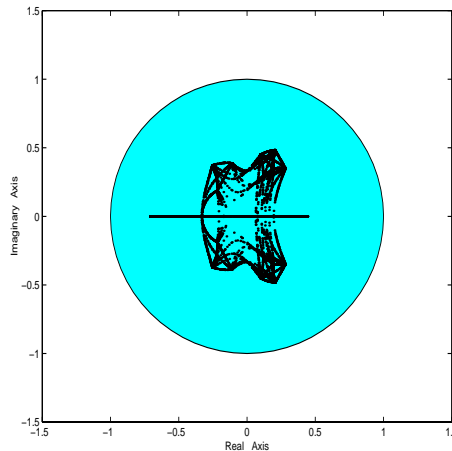


FIG. 4.1. The numerical range of $D(\lambda)$ inscribed in $S(0, 1)$

4.3. Left-half plane and hyperbolic matrix polynomials. To motivate our subsequent analysis, we turn our attention to $n \times n$ quadratic matrix polynomials of the form $P(\lambda) = A_2\lambda^2 + A_1\lambda + A_0$ with real symmetric coefficients such that A_2 is positive definite and A_0, A_1 are positive semidefinite. Such matrix polynomials correspond to second order dynamical models known as *damped vibrating systems* (see [1, 7]). It is not difficult to see that $\sigma(P)$ and $W(P)$ lie in the closed left-half plane of \mathbb{C} . In the following example, we consider a modal approximation to a mechanical damped vibrating system (see [1, pp. 158-161]).

EXAMPLE 4.4. Suppose $P(\lambda)$ is the 2×2 matrix polynomial

$$P(\lambda) = A_2\lambda^2 + A_1\lambda + A_0 = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} \lambda^2 + \begin{pmatrix} 75 & 0 \\ 0 & 15 \end{pmatrix} \lambda + \begin{pmatrix} 256 & 0 \\ 0 & 32 \end{pmatrix}.$$

Its coefficients are real positive definite matrices (and one can verify that for every unit vector $x \in \mathbb{C}^2$, $(x^*A_1x)^2 > 4(x^*A_2x)(x^*A_0x)$). The numerical range $W(P)$ coincides with the union of two closed real intervals, namely, $\mathcal{G}_1 = [-4.6844, -2.5274]$ and $\mathcal{G}_2 = [-45.1207, -7.6674]$, whose endpoints are exactly the four eigenvalues of $P(\lambda)$ [7, Theorem 10.15]. The inner numerical radius of A_2 is $\tilde{r}(A_2) = 0.5858$. Let us first consider the open disk $\Omega = S(-25, 25)$, because it contains $W(P)$ and lies in the open

left-half plane. Let $J = \{0, 1, 2\}$. The bound in Theorem 4.2 is

$$\inf \left\{ \frac{\tilde{r}(A_2) \operatorname{dist}[\lambda, \mathcal{G}_1] \operatorname{dist}[\lambda, \mathcal{G}_2]}{\sqrt{1 + |\lambda|^2 + |\lambda|^4}} : \lambda \in \partial\Omega \right\} = 0.0518$$

(at $\lambda = -50$) and, using Theorem 3.1, $\mathcal{R}_J(P, \Omega) = 0.0631 > 0.0518$. The numerical range $W(P)$, the circle $\partial\Omega$ and the boundaries of the (ϵ, J) -pseudospectra of $P(\lambda)$ for $\epsilon = 0.04, 0.1, 0.2$ are sketched in Figure 4.2. It is apparent that $0.04 < \mathcal{R}_J(P, \Omega) < 0.1$ and that the left-most eigenvalue of $P(\lambda)$, which equals -45.1207 , is quite sensitive to perturbations. This explains the “small” J -stability radius relative to $\Omega = S(-25, 25)$ for this matrix polynomial.

Next, consider Ω to be the open left-half plane. Then for real $t \rightarrow +\infty$, both quantities

$$\frac{1}{\sqrt{1 + t^2 + t^4} \|P(ti)^{-1}\|_2} \quad \text{and} \quad \frac{\tilde{r}(A_2) \operatorname{dist}[ti, \mathcal{G}_1] \operatorname{dist}[ti, \mathcal{G}_2]}{\sqrt{1 + t^2 + t^4}}$$

(here $i = \sqrt{-1}$) are decreasing and converge to $\tilde{r}(A_2)$. (Note that for every $t > 0$, the matrices $P(ti)^{-1}$ and $P(-ti)^{-1} = (P(ti)^*)^{-1}$ have the same norm.) Hence, the J -stability radius of $P(\lambda)$ relative to the open left-half plane is

$$\mathcal{R}_J(P, \Omega) = \lim_{t \rightarrow +\infty} \frac{\tilde{r}(A_2) \operatorname{dist}[ti, \mathcal{G}_1] \operatorname{dist}[ti, \mathcal{G}_2]}{\sqrt{1 + t^2 + t^4}} = \tilde{r}(A_2) = 0.5858,$$

i.e., it coincides with its lower bound in Theorem 4.2.

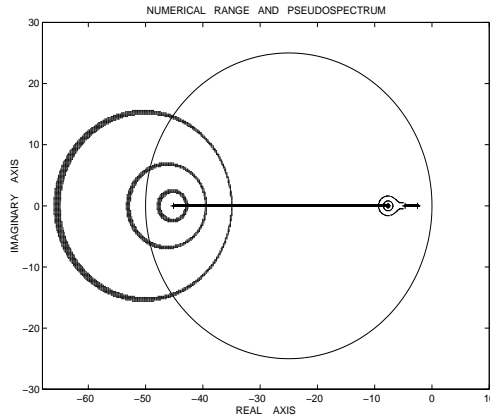


FIG. 4.2. A damped vibrating system.

Motivated by the previous example, we obtain the following more general result.

THEOREM 4.5. *Let $P(\lambda) = A_m \lambda^m + \dots + A_1 \lambda + A_0$ be an $n \times n$ matrix polynomial with bounded numerical range $W(P)$ that lies in the open left-half plane Ω . Suppose that $W(P)$ consists of s connected components $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_s$, and that $\tilde{r}(A_m) = s_{\min}(A_m)$, where $s_{\min}(A_m)$ is the minimum singular value of A_m . Suppose that $m \in J \subseteq \{0, 1, \dots, m\}$. If*

the function

$$g(t) = \frac{\prod_{j=1}^s \text{dist}[ti, \mathcal{G}_j]^{c(\mathcal{G}_j)}}{\sqrt{\sum_{k \in J} t^{2k}}}, \quad t \in \mathbb{R}$$

is increasing in $(-\infty, \mu_1]$ and decreasing in $[\mu_2, \infty)$ for some $\mu_1 < \mu_2$, then $\mathcal{R}_J(P, \Omega)$ equals the minimum of

$$s_{\min}(A_m) \quad \text{and} \quad \inf_{t \in (\mu_1, \mu_2)} \left\{ \frac{1}{\sqrt{\sum_{k \in J} t^{2k}} \|P(ti)^{-1}\|_2} \right\}.$$

Proof. Since $W(P)$ is bounded, $\det A_m \neq 0$ and $P(\lambda)$ has nm finite eigenvalues in Ω . The singular values of the matrix $P(ti)$ are continuous functions of t (see e.g., [20]) and, as $t \rightarrow \pm\infty$, they have asymptotically the behaviour of the singular values of $A_m(ti)^m$. Consequently, $\|P(ti)^{-1}\|_2 = s_{\min}(P(ti))^{-1}$ has asymptotically the behaviour of $\|(A_m t^m)^{-1}\|_2 = |t^{-m}| s_{\min}(A_m)^{-1}$. Also, since $c(\mathcal{G}_1) + c(\mathcal{G}_2) + \dots + c(\mathcal{G}_s) = m$, as $t \rightarrow \pm\infty$, the product $\prod_{j=1}^s \text{dist}[ti, \mathcal{G}_j]^{c(\mathcal{G}_j)}$ behaves asymptotically as $|t^m|$. Thus $\lim_{t \rightarrow \pm\infty} g(t) = 1$ and since $m \in J$,

$$\lim_{t \rightarrow \pm\infty} \frac{1}{\sqrt{\sum_{k \in J} t^{2k}} \|P(ti)^{-1}\|_2} = \lim_{t \rightarrow \pm\infty} (g(t) s_{\min}(A_m)) = s_{\min}(A_m).$$

By the hypothesis that $\tilde{r}(A_m) = s_{\min}(A_m)$ and Theorem 4.1, we have that for every $t \in \mathbb{R}$,

$$\frac{1}{\sqrt{\sum_{k \in J} t^{2k}} \|P(ti)^{-1}\|_2} \geq g(t) s_{\min}(A_m).$$

Hence, by the assumed monotonicity of $g(t)$,

$$\begin{aligned} & \inf \left\{ \frac{1}{\sqrt{\sum_{k \in J} t^{2k}} \|P(ti)^{-1}\|_2} : t \in \mathbb{R} \setminus (\mu_1, \mu_2) \right\} \\ &= \lim_{t \rightarrow \pm\infty} (g(t) s_{\min}(A_m)) = s_{\min}(A_m), \end{aligned}$$

completing the proof. \square

Note that the condition $\tilde{r}(A_m) = s_{\min}(A_m)$ in the above theorem holds when A_m is normal. Based on the proof of the above theorem, when $m \notin J$, the following claim can be made.

COROLLARY 4.6. *Let $P(\lambda)$, Ω and J be as in Theorem 4.5. If $m \notin J$, then there exist $\mu_1, \mu_2 \in \mathbb{R}$ with $\mu_1 < \mu_2$ such that*

$$\mathcal{R}_J(P, \Omega) = \inf \left\{ \frac{1}{\sqrt{\sum_{k \in J} t^{2k}} \|P(ti)^{-1}\|_2} : t \in (\mu_1, \mu_2) \right\}.$$

It is also clear from the definition of $g(t)$ in Theorem 4.5 that the shape of $W(P)$ affects the J -stability radius. In this regard, recall from [12] that an $n \times n$ matrix polynomial $P(\lambda) = A_m\lambda + \dots + A_1\lambda + A_0$ with Hermitian coefficients is said to be *weakly hyperbolic* (resp., *hyperbolic*) if A_m is positive definite and all the roots of the scalar polynomial $x^*P(\lambda)x$ are real (resp., real and distinct) for every $x \in \mathbb{C}^n$. Weakly hyperbolic and hyperbolic matrix polynomials arise in many applications and are of special interest. For example, quadratic hyperbolic matrix polynomials correspond to overdamped vibrating systems as the one in Example 4.4. It is worth noting that the numerical range of a weakly hyperbolic (resp., hyperbolic) matrix polynomial $P(\lambda)$ of m -th degree consists of $s \leq m$ (resp., exactly m) real intervals whose endpoints are eigenvalues of $P(\lambda)$ (see [12]). Hence, Theorem 4.5 yields the following.

COROLLARY 4.7. *Let $P(\lambda) = A_m\lambda^m + \dots + A_1\lambda + A_0$ be an $n \times n$ weakly hyperbolic (resp., hyperbolic) matrix polynomial such that*

$$W(P) = \cup_{j=1}^s [-a_j, -b_j],$$

where $s \leq m$ and

$$0 < b_1 \leq a_1 < b_2 \leq a_2 < \dots < b_s \leq a_s.$$

Let $J = \{0, 1, \dots, m\}$, $m \in J$, and Ω be the open left-half plane. If the function

$$g(t) = \frac{\prod_{j=1}^s |t - b_j|^{c([-a_j, -b_j])}}{\sqrt{1 + t^2 + \dots + t^{2m}}}$$

is decreasing for all $t > 0$, then $\mathcal{R}_J(P, \Omega) = \lambda_{\min}(A_m)$.

In relation to Example 4.4, we can also show the following.

COROLLARY 4.8. *Let $P(\lambda) = A_2\lambda^2 + A_1\lambda + A_0$ be an $n \times n$ selfadjoint matrix polynomial such that $W(P) = [-a_1, -b_1] \cup [-a_2, -b_2]$ ($0 < b_2 \leq a_2 < b_1 \leq a_1$). Let $J = \{0, 1, 2\}$, $2 \in J$, and Ω be the open left-half plane. If for every $t > 0$,*

$$(4.3) \quad (b_1^2 + b_2^2 - 1)t^4 + 2(b_1^2b_2^2 - 1)t^2 + b_1^2b_2^2 - (b_1^2 + b_2^2) > 0,$$

then $\mathcal{R}_J(P, \Omega)$ equals the minimum absolute value amongst the eigenvalues of A_2 .

Proof. Since (4.3) is assumed to hold for every $t > 0$, it is straightforward that the function

$$g(t) = \frac{\sqrt{t^2 + b_1^2} \sqrt{t^2 + b_2^2}}{\sqrt{1 + t^2 + t^4}}$$

is decreasing for all $t > 0$ and increasing for all $t < 0$. The matrix A_2 is Hermitian, yielding $s_{\min}(A_2) = \tilde{r}(A_2)$. Since for every $t > 0$, $\|P(ti)^{-1}\|_2 = \|P(-ti)^{-1}\|_2$, the result follows by applying Theorem 4.5. \square

Notice that (4.3) is always true when $b_1 > 1$ and $b_2 > 1$.

REMARK 4.9. The computation of the stability radius of $A \in \mathcal{M}_n(\mathbb{C})$ relative to the open left-half plane Ω reduces to the computation of $\inf\{s_{\min}(A - tiI) : t \in \mathbb{R}\}$. The existing algorithms in the literature take advantage of the fact that $\lim_{t \rightarrow \pm\infty} s_{\min}(A - tiI) =$

$+\infty$ (see [3, 9]). In contrast, for the J -stability radius of a matrix polynomial $P(\lambda)$, if perturbations of the leading coefficient are allowed, it may be the case that

$$\lim_{t \rightarrow \pm\infty} \frac{1}{\sum_{k \in J} t^{2k} \|P(ti)^{-1}\|_2} < +\infty$$

(see Theorem 4.5). Note also the possible complication that can arise if a perturbation results into a singular leading coefficient. Essentially, this complication gave rise to Corollaries 4.7 and 4.8. On the other hand, if we assume that the leading coefficient of $P(\lambda)$ in (2.1) remains unperturbed ($m \notin J$), then, as in Corollary 4.6,

$$\lim_{t \rightarrow \pm\infty} \frac{1}{\sum_{k \in J} t^{2k} \|P(ti)^{-1}\|_2} = \lim_{t \rightarrow \pm\infty} \frac{\prod_{j=1}^s \text{dist}[ti, \mathcal{G}_j]^{c(\mathcal{G}_j)}}{\sqrt{\sum_{k \in J} t^{2k}}} = +\infty,$$

that is, we have a situation similar to that of matrices.

REFERENCES

- [1] R. E. D. Bishop, G. M. L. Gladwell and S. Michaelson, *Matrix Analysis of Vibration*, Cambridge University Press, Cambridge, 1965.
- [2] A. Böttcher, S. Grudsky and S. Silvermann, Norms of inverses, spectra, and pseudospectra of large truncated Wiener-Hopf operators and Toeplitz matrices, *New York Journal of Mathematics*, **3** (1997), pp. 1-31.
- [3] R. Byers, A bisection method for measuring the distance of a stable matrix to the unstable matrices, *SIAM Journal on Scientific and Statistical Computing*, **9** (1988), pp. 875-881.
- [4] R. Byers and N. K. Nichols, On the stability radius of a generalized state-space system, *Linear Algebra and Its Applications*, **188-189** (1993), pp. 113-134.
- [5] Y. Genin and P. M. Van Dooren, Stability radii of polynomial matrices, *Proceedings of the International Symposium on Computer Aided Control System Design*, pp. 81-84, 1999.
- [6] I. Gohberg, S. Goldberg and M. A. Kaashoek, *Classes of Linear Operators*, Vol. 1, Birkhäuser Verlag, Basel, 1990.
- [7] I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials*, Academic Press, New York, 1982.
- [8] D. Hinrichsen and A. J. Pritchard, Stability radius for structured perturbations and the algebraic Riccati equation, *Systems and Control Letters*, **8** (1986), pp. 105-113.
- [9] C. He and G. A. Watson, An algorithm for computing the distance to instability, *SIAM Journal on Matrix Analysis and Applications*, **20** (1998), pp. 101-116.
- [10] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.
- [11] C.-K. Li and L. Rodman, Numerical range of matrix polynomials, *SIAM Journal on Matrix Analysis and Applications*, **15** (1994), pp. 1256-1265.
- [12] A. S. Markus, *Introduction to the Spectral Theory of Polynomial Operator Pencils*, Transl. of Math. Monographs, Vol. 71, Amer. Math. Society, 1988.
- [13] A. S. Markus and V. I. Matsaev, Some estimates for the resolvent and for the lengths of Jordan chains of an analytic operator function, *Preprint*.
- [14] J. Maroulas and P. Psarrakos, The boundary of numerical range of matrix polynomials, *Linear Algebra and Its Applications*, **267** (1997), pp. 101-111.
- [15] H. Nakazato and P. Psarrakos, On the shape of the numerical range of matrix polynomials, *Linear Algebra and Its Applications*, to appear.
- [16] G. Pappas and D. Hinrichsen, Robust stability of linear systems described by higher order dynamic equations, *IEEE Transactions on Automatic Control*, **38(9)** (1993), pp. 1430-1435.
- [17] P. Psarrakos and M. Tsatsomeris, On the relation between the numerical range and the joint numerical range of matrix polynomials, *Electronic Journal of Linear Algebra*, **6** (2000), pp. 20-30.
- [18] V. Simoncini and E. Gallopoulos, Transfer functions and resolvent norm approximation of large matrices, *Electronic Transactions on Numerical Analysis*, **7** (1998), pp. 190-201.
- [19] J. Sreedhar, P. Van Dooren and A. L. Tits, A fast algorithm to compute the real structured stability radius, *Proceedings of the Conference on Centennial Hurwitz on Stability Theory, Ticino, Switzerland*, 1995.

- [20] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.
- [21] F. Tisseur and N. Higham, Structured pseudospectra for polynomial eigenvalue problems, with applications, *SIAM Journal on Matrix Analysis and Applications*, **23** (1994), pp. 187-208.
- [22] L. N. Trefethen, Pseudospectra of linear operators, *SIAM Review*, **39** (1997), pp. 383-406.
- [23] P. Van Dooren and V. Vermaut, On stability radii of generalized eigenvalue problems, *European Control Conference*, paper FR-M-H6, 1997.