ON THE PSEUDOSPECTRA OF MATRIX POLYNOMIALS

PETER LANCASTER* AND PANAYIOTIS PSARRAKOS†

Abstract. The pseudospectra of a matrix polynomial $P(\lambda)$ are sets of complex numbers that are eigenvalues of matrix polynomials which are near to $P(\lambda)$, i.e., their coefficients are within some fixed magnitude of the coefficients of $P(\lambda)$. Pseudospectra provide important insights into the sensitivity of eigenvalues under perturbations, and have several applications. First, qualitative properties concerning boundedness and connected components of pseudospectra are obtained. Then an accurate continuation algorithm for the numerical determination of the boundary of pseudospectra of matrix polynomials is devised and illustrated. This algorithm is based on a prediction-correction scheme.

Key words. matrix polynomial, eigenvalue, singular value, perturbation, ε-pseudospectrum, boundary, stability

AMS subject classifications. 65F15, 65F35, 93D09

1. Introduction and definitions. Let $\mathbb{C}^{n\times n}$ be the algebra of all $n\times n$ complex matrices and consider the *matrix polynomial*

$$P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_1\lambda + A_0, \tag{1.1}$$

where λ is a complex variable and $A_j \in \mathbb{C}^{n\times n}$, $j = 0, 1, \ldots, m$, with $\det A_m \neq 0$. The study of matrix (and operator) polynomials has a long history, especially with regard to their spectral analysis, see [7, 8, 13, 16]. If $A_j^* = A_j$ $(j = 0, 1, \ldots, m)$, i.e., if all the coefficients of $P(\lambda)$ are hermitian, then $P(\lambda)$ is said to be a *selfadjoint* matrix polynomial.

A scalar $\lambda_0 \in \mathbb{C}$ is said to be an *eigenvalue* of the matrix polynomial $P(\lambda)$ in (1.1) if the system $P(\lambda_0)v = 0$ has a nonzero solution $v_0 \in \mathbb{C}^n$. This solution v_0 is known as an *eigenvector* of $P(\lambda)$ corresponding to λ_0. The set of all eigenvalues of $P(\lambda)$ is the *spectrum* of $P(\lambda)$, namely, $\sigma(P) = \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\}$. Since $\det A_m \neq 0$, $\sigma(P)$ contains no more than nm distinct eigenvalues.

Consider the spectrum of perturbations of the matrix polynomial $P(\lambda)$ in (1.1) of the form

$$P_\Delta(\lambda) = (A_m + \Delta_m)\lambda^m + (A_{m-1} + \Delta_{m-1})\lambda^{m-1} + \cdots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0,$$

where the matrices $\Delta_0, \Delta_1, \ldots, \Delta_m \in \mathbb{C}^{n\times n}$ are arbitrary. A *weighted* pseudospectrum (introduced by Tisseur and Higham [22]) is defined as follows: For a given $\varepsilon > 0$ and a given set of nonnegative weights $\mathbf{w} = \{w_0, w_1, \ldots, w_m\}$ (written $\mathbf{w} \geq 0$), the *ε-pseudospectrum* of $P(\lambda)$ with respect to $\mathbf{w}$ is defined to be

$$\sigma_{\varepsilon,\mathbf{w}}(P) = \{\lambda \in \mathbb{C} : \det P_\Delta(\lambda) = 0,\ \|\Delta_j\| \leq \varepsilon\, w_j,\ j = 0, 1, \ldots, m\},$$

* Department of Mathematics and Statistics, University of Calgary, Calgary AB T2N 1N4, Canada (E-mail: lancaste@math.ucalgary.ca).

†Department of Mathematics, National Technical University, Zografou Campus, 15780 Athens, Greece (E-mail: ppsarr@math.ntua.gr).

where $\|\cdot\|$ is any subordinate matrix norm. The parameters $w_0, w_1, \ldots, w_m \geq 0$ allow freedom in how perturbations are measured; for example, in an absolute sense when $w_0 = w_1 = \cdots = w_m = 1$, or in a relative sense when $w_j = \|A_j\|$ $(j = 0, 1, \ldots, m)$. Also, different values for the w_j admit different levels of confidence in the coefficients A_j. Note also that, when $\varepsilon = 0$, $\sigma_{0,\mathbf{w}}(P) = \sigma(P)$.

Defining the associated compact set of perturbations of $P(\lambda)$,

$$\mathcal{B}(P, \varepsilon, \mathbf{w}) \;=\; \{P_\Delta(\lambda) : \|\Delta_j\| \leq \varepsilon\, w_j,\; j = 0, 1, \ldots, m\},$$

the ε-pseudospectrum of $P(\lambda)$ can also be expressed in the form

$$\sigma_{\varepsilon,\mathbf{w}}(P) \;=\; \{\lambda \in \mathbb{C} : \det P_\Delta(\lambda) = 0,\; P_\Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathbf{w})\}\,.$$

Observe that, if $P(\lambda) = I\lambda - A$ for some $A \in \mathbb{C}^{n\times n}$, then $\sigma(P)$ coincides with the spectrum of A, $\sigma(A)$, in the usual sense. Furthermore, if we set $\mathbf{w} = \{w_0, w_1\} = \{1, 0\}$, then $\sigma_{\varepsilon,\mathbf{w}}(P)$ coincides with the *ε-pseudospectrum* of the matrix A, that is,

$$\sigma_\varepsilon(A) \;=\; \{\lambda \in \mathbb{C} : \lambda \in \sigma(A + \Delta_0),\; \|\Delta_0\| \leq \varepsilon\}\,.$$

The literature on pseudospectra of matrices (and operators) and their applications is extensive, see [2, 4, 15, 25] and the references therein. Also, the special case $n = 1$, is well-understood. Thus, the ε-pseudospectrum of the scalar polynomial $p(\lambda) = a_m\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0$ coincides with the root neighborhood of $p(\lambda)$ introduced by Mosier [20] and the ε-pseudozero set of $p(\lambda)$ investigated by Toh and Trefethen [23].

The case for further development of algorithms for matrix polynomials rests largely on the pervasive second (and higher) degree polynomials used in the analysis of vibrating systems. Efficiencies are gained by avoiding linearizations and, as with problem areas already developed, pseudospectra give valuable insights into the sensitivities of spectra and, particularly, can be expected to clarify the effects of clustered and multiple eigenvalues.

The following lemma is one of the main tools used in this paper. Here, $s_{\min}(\cdot)$ denotes the minimum singular value of a complex matrix. The *spectral norm* is the matrix norm subordinate to the Euclidean vector norm and is consistently used throughout the remainder of this work. First define the scalar polynomial

$$q_{\mathbf{w}}(\lambda) \;=\; w_m\lambda^m + w_{m-1}\lambda^{m-1} + \cdots + w_1\lambda + w_0.$$

LEMMA 1.1. *If the pseudospectrum is defined in terms of the spectral norm then*

$$\sigma_{\varepsilon,\mathbf{w}}(P) \;=\; \{\lambda \in \mathbb{C} : s_{\min}(P(\lambda)) \leq \varepsilon\, q_{\mathbf{w}}(|\lambda|)\}\,. \tag{1.2}$$

Proof. This is just a special case of the important Lemma 2.1 of [22], which applies for any subordinate matrix norm. Here, one needs only that the spectral norm is unitarily invariant. □

As the eigenvalues of $P_\Delta(\lambda)$ are continuous with respect to the entries of the coefficient matrices, it follows from the lemma that the boundary of the ε-pseudospectrum can be written in the form

$$\partial\sigma_{\varepsilon,\mathbf{w}}(P) \;=\; \{\lambda \in \mathbb{C} : s_{\min}(P(\lambda)) = \varepsilon\, q_{\mathbf{w}}(|\lambda|)\}\,, \tag{1.3}$$

and $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$ depends continuously on ϵ.

By using efficient algorithms for computing singular values, equations (1.2) and (1.3) become the main tools for estimation of the ε-pseudospectrum of a matrix polynomial. In particular, Tisseur and Higham (see [22]) obtain a graphical representation of $\sigma_{\varepsilon,\mathbf{w}}(P)$ by evaluating $s_{\min}(P(z))$ on a grid of points in the complex plane. One of our main objectives is the design of an alternative algorithm using a curve-tracing technique.

First of all, however, the geometry and the connected components of pseudospectra of matrix polynomials are studied. In the next two sections, basic boundedness properties of $\sigma_{\varepsilon,\mathbf{w}}(P)$ are obtained, as well as eigenvalue inclusion properties of the connected components of $\sigma_{\varepsilon,\mathbf{w}}(P)$. Section 4 contains the development of a curve-tracing algorithm for computing the boundary of pseudospectra. Our objective is to demonstrate that curve-following procedures for graphing pseudospectra for the classical eigenvalue problem (see [1, 2]) can be extended to apply directly to matrix polynomials with nonsingular leading coefficients. Numerical examples are given in the last section to demonstrate the feasibility of the method. Optimization of the algorithm with respect to efficiency is not considered here. Generalization of the method to admit matrix polynomials with singular leading coefficients (using ideas developed in [9]) should also be possible.

2. General properties. If the matrix polynomial $P(\lambda)$ in (1.1) is real (i.e., all its coefficients are real matrices) or selfadjoint, then it is well known that the spectrum of $P(\lambda)$ is symmetric with respect to the real axis. This symmetry also holds for the pseudospectra of matrix polynomials.

PROPOSITION 2.1. *Let $P(\lambda) = A_m\lambda^m + \cdots + A_1\lambda + A_0$ be an $n \times n$ real or selfadjoint matrix polynomial. Then for any $\varepsilon > 0$ and $\mathbf{w} \geq 0$, the ε-pseudospectrum $\sigma_{\varepsilon,\mathbf{w}}(P)$ is symmetric with respect to the real axis.*

Proof. Suppose $\mu \in \sigma_{\varepsilon,\mathbf{w}}(P)$. Then there is a matrix polynomial $(A_m+\Delta_m)\lambda^m + \cdots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0 \in \mathcal{B}(P, \varepsilon, \mathbf{w})$ such that

$$\det\left[(A_m + \Delta_m)\mu^m + \cdots + (A_1 + \Delta_1)\mu + A_0 + \Delta_0\right] \;=\; 0.$$

If the coefficients of $P(\lambda)$ are real, then

$$\det\left[\overline{(A_m + \Delta_m)\mu^m + \cdots + (A_1 + \Delta_1)\mu + A_0 + \Delta_0}\right] \;=$$

$$\det\left[(A_m + \overline{\Delta}_m)\overline{\mu}^m + \cdots + (A_1 + \overline{\Delta}_1)\overline{\mu} + A_0 + \overline{\Delta}_0\right] \;=\; 0,$$

where $(A_m + \overline{\Delta}_m)\lambda^m + \cdots + (A_1 + \overline{\Delta}_1)\lambda + A_0 + \overline{\Delta}_0$ also lies in $\mathcal{B}(P, \varepsilon, \mathbf{w})$. Hence, $\overline{\mu} \in \sigma_{\varepsilon,\mathbf{w}}(P)$.

Similarly, if all the coefficients of $P(\lambda)$ are hermitian, then

$$\det\left[(A_m+\Delta_m)\mu^m+\cdots+(A_1+\Delta_1)\mu+A_0+\Delta_0\right]^* \;=$$

$$\det\left[(A_m+\Delta_m^*)\overline{\mu}^m+\cdots+(A_1+\Delta_1^*)\overline{\mu}+A_0+\Delta_0^*\right] \;=\; 0,$$

where $(A_m+\Delta_m^*)\lambda^m+\cdots+(A_1+\Delta_1^*)\lambda+A_0+\Delta_0^* \in \mathcal{B}(P,\varepsilon,\mathbf{w})$. Thus, $\overline{\mu}$ lies in $\sigma_{\varepsilon,\mathbf{w}}(P)$. □

Notice that, if there is a perturbation $P_\Delta(\lambda)\in\mathcal{B}(P,\varepsilon,\mathbf{w})$ with identically zero determinant, then $\sigma_{\varepsilon,\mathbf{w}}(P)$ coincides with the whole complex plane so, *a priori*, the pseudospectrum may be unbounded. On the other hand, since the leading coefficient A_m is nonsingular, the matrix polynomial $P(\lambda)$ has exactly nm (finite) eigenvalues, counting multiplicities, so that for ε sufficiently small, the pseudospectrum must be bounded and consist of no more than nm connected components. By extending a technique of Li and Rodman [14, Theorem 2.3], the next result establishes a necessary and sufficient condition for $\sigma_{\varepsilon,\mathbf{w}}(P)$ to be bounded (see also Proposition 3.5 of [10] for a similar result on stability radii of matrix polynomials).

THEOREM 2.2. *Let $P(\lambda)$ be an $n\times n$ matrix polynomial as in (1.1). Then the pseudospectrum $\sigma_{\varepsilon,\mathbf{w}}(P)$ is bounded if and only if the εw_m-pseudospectrum of the leading coefficient A_m of $P(\lambda)$ does not contain the origin.*

Proof. For fixed $\varepsilon>0$ and $\mathbf{w}=\{w_0,w_1,\ldots,w_m\}\geq 0$, suppose that $0\notin\sigma_{\varepsilon w_m}(A_m)$. Then $\det(A_m+\Delta_m)\neq 0$ whenever $\|\Delta_m\|\leq\varepsilon\, w_m$, and

$$\zeta_\varepsilon \;=\; \min\{|\det(A_m+\Delta_m)| : \|\Delta_m\|\leq\varepsilon\, w_m\} \;>\; 0.$$

Since the set $\mathcal{B}(P,\varepsilon,\mathbf{w})$ is compact, there is an $M_\varepsilon>0$ such that for any perturbation of $P(\lambda)$,

$$P_\Delta(\lambda) \;=\; (A_m+\Delta_m)\lambda^m+\cdots+(A_1+\Delta_1)\lambda+A_0+\Delta_0 \;\in\; \mathcal{B}(P,\varepsilon,\mathbf{w}),$$

and for any $\lambda\in\mathbb{C}$ with $|\lambda|>M_\varepsilon$, we have

$$|\det P_\Delta(\lambda)-\det(A_m+\Delta_m)\lambda^{nm}| \;<\; \zeta_\varepsilon|\lambda^{mn}| \;\leq\; |\det(A_m+\Delta_m)\lambda^{mn}|$$

(keeping in mind that $\det P_\Delta(\lambda)$ is a scalar polynomial with leading term $\det(A_m+\Delta_m)\lambda^{mn}$). Hence, $\det P_\Delta(\lambda)\neq 0$, i.e., $\sigma_{\varepsilon,\mathbf{w}}(P)\subseteq\{\lambda\in\mathbb{C}:|\lambda|\leq M_\varepsilon\}$; $\sigma_{\varepsilon,\mathbf{w}}(P)$ is bounded.

To prove the converse, assume that $\sigma_{\varepsilon,\mathbf{w}}(P)$ is bounded but there is a

$$P_{\hat\Delta}(\lambda) \;=\; (A_m+\hat\Delta_m)\lambda^m+\cdots+(A_1+\hat\Delta_1)\lambda+A_0+\hat\Delta_0 \;\in\; \mathcal{B}(P,\varepsilon,\mathbf{w})$$

with $\det(A_m+\hat\Delta_m)=0$. Then at least one of the coefficients of the scalar polynomial $\det P_{\hat\Delta}(\lambda)$ is nonzero; otherwise $\sigma_{\varepsilon,\mathbf{w}}(P)=\mathbb{C}$, a contradiction. Let the coefficient of λ^τ ($\tau\in\{0,1,\ldots,nm-1\}$) in $\det P_{\hat\Delta}(\lambda)$ be nonzero and denote it by β_τ. Construct a sequence $\{\hat\Delta_{m,k}\}_{k\in\mathbb{N}}\subset\mathbb{C}^{n\times n}$ such that $\lim_{k\to\infty}\hat\Delta_{m,k}=\hat\Delta_m$ and

$$\det(A_m+\hat\Delta_{m,k})\neq 0 \;\;\text{ and }\;\; \|\hat\Delta_{m,k}\|\leq\varepsilon\, w_m \quad (k=1,2,\ldots).$$

Clearly, for a fixed $\delta > 0$, $|\det(A_m + \hat{\Delta}_{m,k})| < \delta$ for all sufficiently large k. Since $\sigma_{\varepsilon,\mathbf{w}}(P)$ is bounded, the $(nm-\tau)$th elementary symmetric function of the roots of $\det[(A_m + \hat{\Delta}_{m,k})\lambda^m + \cdots + (A_1 + \hat{\Delta}_1)\lambda + A_0 + \hat{\Delta}_0]$, which is equal to $\pm\,\beta_\tau / \det(A_m + \hat{\Delta}_{m,k})$, is bounded for all k. This contradicts the construction of the sequence $\{\hat{\Delta}_{m,k}\}_{k\in\mathbb{N}}$. □

Theorem 2 of [20] can be generalized as follows:

THEOREM 2.3. *If $\sigma_{\varepsilon,\mathbf{w}}(P)$ is bounded, then it has no more than nm connected components, and any $P_\Delta(\lambda) \in \mathcal{B}(P,\varepsilon,\mathbf{w})$ has an eigenvalue in each one of these components. Furthermore, $P(\lambda)$ and $P_\Delta(\lambda)$ have the same number of eigenvalues, counting multiplicities, in each connected component of $\sigma_{\varepsilon,\mathbf{w}}(P)$.*

Proof. Suppose $\sigma_{\varepsilon,\mathbf{w}}(P)$ is bounded. It follows from Theorem 2.2 that, for any perturbation of $P(\lambda)$,

$$P_\Delta(\lambda) \;=\; (A_m + \Delta_m)\lambda^m + \cdots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0 \;\in\; \mathcal{B}(P,\varepsilon,\mathbf{w}),$$

$\det(A_m + \Delta_m) \neq 0$, i.e., the leading coefficient of the polynomial $\det P_\Delta(\lambda)$ is nonsingular. As a consequence, $P_\Delta(\lambda)$ has exactly nm eigenvalues, counting multiplicities, as does every member of the family of matrix polynomials

$$P_{\Delta,t}(\lambda) \;=\; (1-t)\,P(\lambda) + tP_\Delta(\lambda) \;\;;\quad t \in [0,1].$$

Moreover, for any $t \in [0,1]$, $P_{\Delta,t}(\lambda)$ belongs to $\mathcal{B}(P,\varepsilon,\mathbf{w})$ and its eigenvalues lie in $\sigma_{\varepsilon,\mathbf{w}}(P)$.

The coefficients of the scalar polynomial $\det P_{\Delta,t}(\lambda)$ are continuous functions of $t \in [0,1]$. Hence, by the continuity of the zeros of $\det P_{\Delta,t}(\lambda)$ with respect to its coefficients, as t varies from 0 to 1, the eigenvalues of $P_{\Delta,t}(\lambda)$ trace continuous paths from the eigenvalues of $P(\lambda)$ $(= P_{\Delta,0}(\lambda))$ to the eigenvalues of $P_\Delta(\lambda)$ $(= P_{\Delta,1}(\lambda))$. Thus, if $P(\lambda)$ has k eigenvalues (counting multiplicities) in a connected component $\mathcal{G}$ of $\sigma_{\varepsilon,\mathbf{w}}(P)$ and its $nm-k$ remaining eigenvalues are isolated in $\sigma_{\varepsilon,\mathbf{w}}(P) \setminus \mathcal{G}$, then this is true for the eigenvalues of every $P_{\Delta,t}(\lambda)$, $t \in [0,1]$. Consequently, $P_\Delta(\lambda)$ has exactly k eigenvalues in $\mathcal{G}$, counting multiplicities.

Finally, note that each bounded connected component of $\sigma_{\varepsilon,\mathbf{w}}(P)$ contains at least one eigenvalue of $P(\lambda)$, and by the above discussion, it contains at least one eigenvalue of the perturbation $P_\Delta(\lambda)$. Hence, $\sigma_{\varepsilon,\mathbf{w}}(P)$ cannot have more than nm connected components. □

Since the origin lies in $\sigma_{\varepsilon w_m}(A_m)$ if and only if $s_{\min}(A_m) \leq \varepsilon\, w_m$, it also follows that:

COROLLARY 2.4. *For any $\varepsilon > 0$ such that $\varepsilon\, w_m < s_{\min}(A_m)$, $\sigma_{\varepsilon,\mathbf{w}}(P)$ consists of no more than nm bounded connected components.*

3. Matrix polynomials with bounded numerical range. The *numerical range* of the matrix polynomial $P(\lambda)$ is defined by

$$W(P) \;=\; \{\lambda \in \mathbb{C} : v^*P(\lambda)v = 0,\, v \in \mathbb{C}^n,\, v^*v = 1\}$$

(see e.g., [14, 17, 18]), and it is always closed and contains $\sigma(P)$. For the linear pencil $I\lambda - A$ $(A \in \mathbb{C}^{n\times n})$, $W(I\lambda - A)$ coincides with the classical *numerical range* (also

known as the *field of values*) of the matrix A, $F(A) = \{v^*Av : v \in \mathbb{C}^n,\ v^*v = 1\}$, which is always compact and convex [11]. The *inner numerical radius* of A is defined by $\hat{r}(A) = \min\{|\lambda| : \lambda \in \partial F(A)\}$. By Theorem V.3.2 of [12], the ε-pseudospectrum of A lies in the region

$$F(A) + S(0,\varepsilon) \;=\; \{\lambda \in \mathbb{C} : \mathrm{dist}(\lambda, F(A)) \le \varepsilon\},$$

where $\mathrm{dist}(\lambda, F(A))$ denotes the distance between the point λ and the numerical range $F(A)$. Thus, $F(A) + S(0,\varepsilon)$ is often used as an initial region for the estimation of $\sigma_\varepsilon(A)$.

The numerical range $W(P)$ is bounded if and only if $0 \notin F(A_m)$, and in this case, it has no more than m connected components [14]. If $\mathcal{G}$ is a bounded connected component of $W(P)$, then for any unit vector $v \in \mathbb{C}^n$, the number of zeros of the scalar polynomial $v^*P(\lambda)v$ in $\mathcal{G}$, counting multiplicities, does not depend on v [16, Lemma 26.8], i.e., it is constant. If we denote this constant number by $c(\mathcal{G})$, then $P(\lambda)$ has exactly $n\,c(\mathcal{G})$ eigenvalues in $\mathcal{G}$, counting multiplicities, [18, Theorem 2.1]. Now Theorem V.3.2 of [12] can be generalized to the case of matrix polynomials. Although this result is mainly of theoretical interest, note that estimates of the inner numerical radius are available in [3].

THEOREM 3.1. *Let $P(\lambda) = A_m\lambda^m + \cdots + A_1\lambda + A_0$ be an $n \times n$ matrix polynomial with bounded numerical range that consists of ξ ($\le m$) connected components $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_\xi$. Then for given $\varepsilon > 0$ and $\mathbf{w} \ge 0$,*

$$\sigma_{\varepsilon,\mathbf{w}}(P) \;\subseteq\; \left\{ \lambda \in \mathbb{C} : \prod_{j=1}^{\xi} \mathrm{dist}(\lambda, \mathcal{G}_j)^{c(\mathcal{G}_j)} \;\le\; \frac{\varepsilon\, q_{\mathbf{w}}(|\lambda|)}{\hat{r}(A_m)} \right\},$$

where it is assumed that $\mathrm{dist}(\lambda, \mathcal{G}_j) = 0$ when $\lambda \in \mathcal{G}_j$.

Proof. Suppose that $\lambda_0 \in \sigma_{\varepsilon,\mathbf{w}}(P)$. Then there exist a perturbation of $P(\lambda)$,

$$P_\Delta(\lambda) \;=\; (A_m + \Delta_m)\lambda^m + \cdots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0 \;\in\; \mathcal{B}(P, \varepsilon, \mathbf{w}),$$

and a unit vector $v_0 \in \mathbb{C}^n$ such that $P_\Delta(\lambda_0)v_0 = 0$. Hence,

$$v_0^*(\Delta_m\lambda_0^m + \cdots + \Delta_1\lambda_0 + \Delta_0)v_0 \;=\; -v_0^*P(\lambda_0)v_0,$$

and consequently,

$$\sum_{j=0}^{m} |v_0^*\Delta_j v_0|\,|\lambda_0|^j \ge \left| \sum_{j=0}^{m} (v_0^*\Delta_j v_0)\,\lambda_0^j \right| \;=\; |v_0^*P(\lambda_0)v_0|$$
$$= |v_0^*A_m v_0| \prod_{j=1}^{m} |\lambda_0 - \lambda_j(v_0)|,$$

where $\lambda_1(v_0), \lambda_2(v_0), \ldots, \lambda_m(v_0)$ are the zeros of the polynomial $v_0^*P(\lambda)v_0$. Since $\varepsilon\, w_j \ge \|\Delta_j\| \ge |v_0^*\Delta_j v_0|$, $j = 0, 1, \ldots, m$, it follows that

$$\varepsilon\, q_{\mathbf{w}}(|\lambda_0|) \ge \sum_{j=0}^{m} \|\Delta_j\|\,|\lambda_0|^j \;\ge\; \sum_{j=0}^{m} |v_0^*\Delta_j v_0|\,|\lambda_0|^j$$

$$\geq \hat{r}(A_m) \prod_{j=1}^{m} |\lambda_0 - \lambda_j(v_0)| \;\geq\; \hat{r}(A_m) \prod_{j=1}^{\xi} \mathrm{dist}(\lambda_0, \mathcal{G}_j)^{c(\mathcal{G}_j)}.$$

The proof is complete. □

4. A curve-tracing algorithm. For fixed $\varepsilon > 0$ and $\mathbf{w} = \{w_0, w_1, \ldots, w_m\} \geq 0$, the boundary of the ε-pseudospectrum of the matrix polynomial $P(\lambda) = A_m\lambda^m + \cdots + A_1\lambda + A_0$ is given by (1.3). In this section, we describe a prediction-correction continuation methodology for the computation of $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$, extending an algorithm of Brühl [1, 2] for matrices. As noted in Section 1, for our purposes, the pseudospectrum is that defined in terms of the spectral norm.

For convenience, define the function

$$g_P(x, y) \;=\; s_{\min}(P(x + \mathrm{i}y)) \;\;;\;\; x, y \in \mathbb{R}.$$

When there is no fear of confusion, we write $g_P(\lambda) = s_{\min}(P(\lambda))$ for $\lambda \in \mathbb{C}$.

Our approach is based on the following result of Sun [21] concerning the differentiability of simple singular values.

THEOREM 4.1. *Let the matrix valued function $F(\mathbf{b}) : \mathbb{R}^d \mapsto \mathbb{C}^{n\times n}$ be real analytic in a neighborhood of $\mathbf{b}_0 = \left(b_0^{(1)}, b_0^{(2)}, \ldots, b_0^{(d)}\right) \in \mathbb{R}^d$. Suppose that s_0 is a simple nonzero singular value of $F(\mathbf{b}_0)$, and u_0 and v_0 are associated left and right singular vectors, respectively.*

Then there is a neighborhood $\mathcal{N}$ of $\mathbf{b}_0$ on which a simple nonzero singular value $s(\mathbf{b})$ of $F(\mathbf{b})$ is defined together with corresponding left and right singular vectors $u(\mathbf{b})$ and $v(\mathbf{b})$, respectively, such that $s(\mathbf{b}_0) = s_0$, $u(\mathbf{b}_0) = u_0$ and $v(\mathbf{b}_0) = v_0$, and the functions s, u and v are real analytic on $\mathcal{N}$.

Moreover, the partial derivatives of $s(\mathbf{b})$ at $\mathbf{b}_0$ are given by

$$\frac{\partial s(\mathbf{b}_0)}{\partial b^{(j)}} \;=\; \mathrm{Re}\left(u_0^* \,\frac{\partial F(\mathbf{b}_0)}{\partial b^{(j)}}\, v_0\right) \;\;;\;\; j = 1, 2, \ldots, d.$$

The next corollary is a direct consequence of this theorem.

COROLLARY 4.2. *Let $\lambda_0 = x_0 + \mathrm{i}y_0 \in \mathbb{C} \setminus \sigma(P)$. If $s_{\min}(P(\lambda_0))$ is a simple singular value of the matrix $P(\lambda_0)$, and u_0, v_0 are associated left and right singular vectors, respectively, then $g_P(x, y)$ is real analytic in a neighborhood of $(x_0, y_0) \in \mathbb{R}^2$, and*

$$\nabla g_P(x_0, y_0) \;=\; \left(\mathrm{Re}\left(u_0^* \,\frac{\partial P(x_0 + \mathrm{i}y_0)}{\partial x}\, v_0\right), \mathrm{Re}\left(u_0^* \,\frac{\partial P(x_0 + \mathrm{i}y_0)}{\partial y}\, v_0\right)\right).$$

The basic continuation method investigated here for finding points on the implicitly defined curve

$$\partial\sigma_{\varepsilon,\mathbf{w}}(P) \;=\; \{\lambda \in \mathbb{C} : g_P(\lambda) - \varepsilon\, q_{\mathbf{w}}(|\lambda|) = 0\},$$

consists of an initial step to find a starting point on the curve followed by a sequence of "predictor" steps tangential to the boundary and "corrector" steps to go back to the boundary.

Initial step: For calculation of a first point on the boundary of $\sigma_{\varepsilon,\mathbf{w}}(P)$, let $\lambda_0 = x_0 + \mathrm{i}y_0 \in \sigma_{\varepsilon,\mathbf{w}}(P) \setminus \sigma(P)$ and $d_0 \in \mathbb{C}$ be nonzero, and consider the function

$$h(t) \;=\; g_P(\lambda_0 + t\,d_0) - \varepsilon\, q_{\mathbf{w}}(|\lambda_0 + t\,d_0|) \;\;;\quad t \in \mathbb{R}.$$

Then use Newton's method to find a solution of $h(t) = 0$ along the straight line $\{\lambda_0 + t\,d_0 : t \in \mathbb{R}\}$ in the complex plane. Without loss of generality, it may be assumed that the initial value of t is $t_0 = 0$. Moreover, we assume that g_P is differentiable at λ_0 and the gradient $\nabla g_P(x_0, y_0)$ given by Corollary 4.2 is nonzero.

Then the first Newton iterate is

$$t_1 \;=\; t_0 - \frac{h(t_0)}{h'(t_0)} \;=\; -\,\frac{g_P(\lambda_0) - \varepsilon\, q_{\mathbf{w}}(|\lambda_0|)}{h'(0)}\,,$$

where

$$\begin{aligned} h'(0) = \frac{\partial[g_P(\lambda_0) - \varepsilon\, q_{\mathbf{w}}(|\lambda_0|)]}{\partial d_0} \;&=\; \frac{\partial\left[g_P(x_0 + \mathrm{i}y_0) - \varepsilon\, q_{\mathbf{w}}\left(\sqrt{x_0^2 + y_0^2}\right)\right]}{\partial d_0} \\ &= (\mathrm{Re}\, d_0)\left[\mathrm{Re}\left(u_0^*\,\frac{\partial P(x_0 + \mathrm{i}y_0)}{\partial x}\;v_0\right) - \varepsilon\,\frac{\partial q_{\mathbf{w}}\left(\sqrt{x_0^2 + y_0^2}\right)}{\partial x}\right] \\ &\quad + (\mathrm{Im}\, d_0)\left[\mathrm{Re}\left(u_0^*\,\frac{\partial P(x_0 + \mathrm{i}y_0)}{\partial y}\;v_0\right) - \varepsilon\,\frac{\partial q_{\mathbf{w}}\left(\sqrt{x_0^2 + y_0^2}\right)}{\partial y}\right]. \end{aligned}$$

For brevity, write the partial derivatives on the right in the form

$$\mathrm{Re}\left(u_0^*\,\frac{\partial P(x_0 + \mathrm{i}y_0)}{\partial x}\;v_0\right) - \varepsilon\,\frac{\partial q_{\mathbf{w}}\left(\sqrt{x_0^2 + y_0^2}\right)}{\partial x} = \mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, x)$$

and

$$\mathrm{Re}\left(u_0^*\,\frac{\partial P(x_0 + \mathrm{i}y_0)}{\partial y}\;v_0\right) - \varepsilon\,\frac{\partial q_{\mathbf{w}}\left(\sqrt{x_0^2 + y_0^2}\right)}{\partial y} = \mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, y).$$

Then

$$(4.1)\;\; z_1 \;=\; \lambda_0 + t_1 d_0 \;=\; x_0 + \mathrm{i}y_0 - \frac{g_P(x_0 + \mathrm{i}y_0) - \varepsilon\, q_{\mathbf{w}}\left(\sqrt{x_0^2 + y_0^2}\right)}{(\mathrm{Re}\, d_0)\mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, x) + (\mathrm{Im}\, d_0)\mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, y)}\, d_0.$$

Since the point $\lambda_0 = x_0 + \mathrm{i}y_0$ lies in $\sigma_{\varepsilon,\mathbf{w}}(P)$, for suitably chosen direction d_0, we can estimate a boundary point of $\sigma_{\varepsilon,\mathbf{w}}(P)$ (in particular, an intersection point of $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$

and $\{\lambda_0 + t\,d_0 : t \in \mathbb{R}\}$) by repeating (4.1) until the quantity $|s_{\min}(P(z)) - \varepsilon\, q_{\mathbf{w}}(|z|)|$ is small enough.

It is also worth noting that

$$(\operatorname{Re} d_0)\mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, x) + (\operatorname{Im} d_0)\mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, y) \;=\; \operatorname{Re}\left[\,\overline{d}_0\,(\mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, x) + \mathrm{i}\mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, y))\right],$$

and on choosing the direction $d_0 = \mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, x) + \mathrm{i}\mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, y)$, the equation (4.1) implies

$$z_1 \;=\; x_0 + \mathrm{i}y_0 - \frac{g_P(x_0 + \mathrm{i}y_0) - \varepsilon\, q_{\mathbf{w}}\left(\sqrt{x_0^2 + y_0^2}\right)}{\mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, x) - \mathrm{i}\mathcal{R}_{\varepsilon,\mathbf{w}}(\lambda_0, y)}\,. \tag{4.2}$$

Prediction: Assuming now that the point $z_{k-1} \in \partial\sigma_{\varepsilon,\mathbf{w}}(P)$ has been computed and τ_k is the corresponding step-length, the (tangential) prediction for the kth boundary point of $\sigma_{\varepsilon,\mathbf{w}}(P)$, z_k, is

$$\begin{aligned}\hat{z}_k &= z_{k-1} + \tau_k \left(\mathrm{i}\,\frac{\nabla\left[g_P(z_{k-1}) - \varepsilon\, q_{\mathbf{w}}(|z_{k-1}|)\right]}{|\nabla\left[g_P(z_{k-1}) - \varepsilon\, q_{\mathbf{w}}(|z_{k-1}|)\right]|}\right)\\ &= z_{k-1} + \tau_k \left(\mathrm{i}\,\frac{(\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, x), \mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, y))}{|(\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, x), \mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, y))|}\right),\end{aligned}$$

i.e., the direction tangential to $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$ is chosen.

Correction: For sufficiently small τ_k, the correction step consists of a single Newton iterate with respect to the equation $g_P(\hat{z}_k + t\,d_k) - \varepsilon\, q_{\mathbf{w}}(|\hat{z}_k + t\,d_k|) = 0$, with an appropriate direction d_k and initial value $t_0 = 0$. It has been found that one Newton step gives adequate numerical performance, although the effect of taking more steps at this stage could be a subject for further investigation.

A natural choice for d_k is the gradient

$$\nabla\left[g_P(\hat{z}_k) - \varepsilon\, q_{\mathbf{w}}(|\hat{z}_k|)\right] \;=\; (\mathcal{R}_{\varepsilon,\mathbf{w}}(\hat{z}_k, x), \mathcal{R}_{\varepsilon,\mathbf{w}}(\hat{z}_k, y))\,.$$

In this case, the step (4.2) is applied, and the estimation of z_k requires the computation of the singular values $s_{\min}(P(z_{k-1}))$ and $s_{\min}(P(\hat{z}_k))$, and their associated left and right singular vectors.

However, the computation of $s_{\min}(P(\hat{z}_k))$ and the corresponding singular vectors can be avoided (and the computational cost of the algorithm reduced by about a half) if the correction step is taken in the direction of

$$\nabla\left[g_P(z_{k-1}) - \varepsilon\, q_{\mathbf{w}}(|z_{k-1}|)\right] \;=\; (\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, x), \mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, y))$$

and (4.2) is written in the form

$$z_k \;=\; \hat{z}_k - \frac{g_P(z_{k-1}) - \varepsilon\, q_{\mathbf{w}}(|z_{k-1}|)}{\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, x) - \mathrm{i}\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, y)}\,.$$

In Figure 4.1, a graphical illustration of the directions $d_k = \nabla\left[g_P(\hat{z}_k) - \varepsilon\, q_{\mathbf{w}}(|\hat{z}_k|)\right]$ (left part) and $d_k = \nabla\left[g_P(z_{k-1}) - \varepsilon\, q_{\mathbf{w}}(|z_{k-1}|)\right]$ (right part) is given. The prediction points are marked with asterisks, the correction points are plotted as 'o', and ∇ denotes the gradient $\nabla\left[g_P(z_{k-1}) - \varepsilon\, q_{\mathbf{w}}(|z_{k-1}|)\right]$. Note that in the right part of the figure, the line segment $[\hat{z}_k, z_k]$ is parallel to ∇.

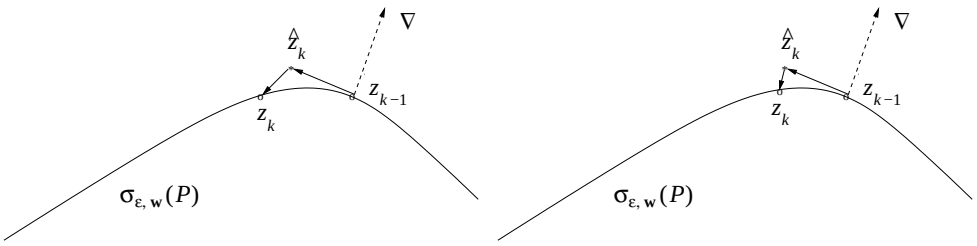

FIG. 4.1. *Choosing the direction d_k in the correction step.*

Algorithm

Input: The coefficients $A_0, A_1, \ldots, A_m$ of the matrix polynomial $P(\lambda)$,
the parameter $\varepsilon > 0$ and the weights $w_0, w_1, \ldots, w_m \geq 0$,
the number N of points to be determined on $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$,
an approximation μ_0 of an eigenvalue of $P(\lambda)$,
the search direction d_0 for the initial point,
the relative accuracy *tol* for the initial point,
the step-length r.

Step 1: *(For computing the first boundary point)*

Set $z_{\text{new}} = \mu_0 + \varepsilon\, d_0$.

While $|s_{\min}(P(z_{\text{new}})) - \varepsilon\, q_{\mathbf{w}}(|z_{\text{new}}|)| > \varepsilon\, q_{\mathbf{w}}(|z_{\text{new}}|)\, tol$, repeat:

(a) Set $z_{\text{old}} = z_{\text{new}}$.
(b) Compute the minimum singular value $s_{\min}$ of $P(z_{\text{old}})$ and associated left and right singular vectors u_0 and v_0.
(c) Compute the gradient

$$\nabla\left[g_P(z_{\text{old}}) - \varepsilon\, q_{\mathbf{w}}(|z_{\text{old}}|)\right] = \left(\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{\text{old}}, x), \mathcal{R}_{\varepsilon,\mathbf{w}}(z_{\text{old}}, y)\right).$$

(d) Compute the next Newton iterate

$$z_{\text{new}} = z_{\text{old}} - \frac{s_{\min}(P(z_{\text{old}})) - \varepsilon\, q_{\mathbf{w}}(|z_{\text{old}}|)}{(\operatorname{Re} d_0)\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{\text{old}}, x) + (\operatorname{Im} d_0)\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{\text{old}}, y)}\, d_0.$$

End while

Set $z_1 = z_{\text{new}}$.

Step 2: *(For computing the remaining $N-1$ boundary points)*

For $k = 2, 3, \ldots, N,$ repeat:

(a) Compute the minimum singular value $s_{\min}$ of $P(z_{k-1})$ and associated left and right singular vectors u_0 and v_0.

(b) Compute the gradient

$$\nabla \left[g_P(z_{k-1}) - \varepsilon\, q_{\mathbf{w}}(|z_{k-1}|) \right] \;=\; (\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, x), \mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, y))$$

and the direction

$$d_k \;=\; \mathrm{i}\, \frac{\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, x) + \mathrm{i}\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, y)}{|\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, x) + \mathrm{i}\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, y)|}\,.$$

(c) Compute the predicted point

$$\hat{z}_k \;=\; z_{k-1} + d_k r.$$

(d) Compute the corrected point

$$z_k \;=\; \hat{z}_k - \frac{s_{\min}(P(z_{k-1})) - \varepsilon\, q_{\mathbf{w}}(|z_{k-1}|)}{\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, x) - \mathrm{i}\mathcal{R}_{\varepsilon,\mathbf{w}}(z_{k-1}, y)}\,.$$

End for

Output: The points $z_1, z_2, \ldots, z_N$.

Of course, this algorithm tracks the boundary of that connected component of the pseudospectrum $\sigma_{\varepsilon,\mathbf{w}}(P)$ containing μ_0. Consequently, for a complete picture of $\sigma_{\varepsilon,\mathbf{w}}(P)$, it may be necessary to repeat the procedure several times with different (appropriate) values of μ_0.

Since the size of the step-lengths τ_k $(k = 2, 3, \ldots, N)$ used in the prediction step affects the accuracy and the computational cost of the algorithm, it is important to obtain criteria for their selection. This is an open problem, which is partially solved by Bekas and Gallopoulos [1]: an efficient choice of τ_k will undoubtedly depend mainly on the local shape of the curve $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$. The present experiments are confined to a constant step-length, r (see (c) of Step 2). In particular, it is apparent that a procedure in which step-size is related to an estimate of the local curvature might be advantageous. This and related questions must, however, be deferred to future research.

As might be expected, difficulties appear near points of $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$ where the function g_P is not differentiable and the minimum singular value is multiple, and also near points where the distance between distinct connected components of $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$ becomes relatively small (see Example 5.2 below). In these cases, the curve-tracing algorithm may lose its path or retrace its own steps. Some of these difficulties can be resolved by choosing a smaller r (increasing the cost) or/and more suitable values for μ_0 and d_0.

In spite of these apparent weaknesses, the authors' experiments suggest that the algorithm is remarkably robust.

An important feature of this technique is that it does not require *a priori* knowledge of the size or shape of $\sigma_{\varepsilon,\mathbf{w}}(P)$, since it determines the connected components of the pseudospectrum one after the other by using starting points close to eigenvalues. Moreover, the cost does not depend strongly on the degree of the matrix polynomial $P(\lambda)$. This parameter, m, appears only in the calculation of $P(z)$, $q(|z|)$ and their partial derivatives. The main cost of the algorithm comes from the computation of the singular values $s_{\min}(P(z_{k-1}))$, $k = 2, 3, \ldots, N$, and associated singular vectors. For this task, the suggested reference is [22], where five techniques (the Golub-Reinsch SVD algorithm, the transfer function approach and the computation of solvents by using the generalized Schur decomposition or the Newton method or the Bernoulli iteration) are compared in terms of flops and execution time (see also [15, 24]). In the examples of the next section, the SVD technique has been used.

5. Numerical examples. We give some examples to illustrate our results and the use of the proposed algorithm.

EXAMPLE 5.1. (A wing problem) Consider the quadratic matrix polynomial $Q(\lambda) = A_2\lambda^2 + A_1\lambda + A_0$ with

$$A_2 = \begin{bmatrix} 17.6 & 1.28 & 2.89 \\ 1.28 & 0.824 & 0.413 \\ 2.89 & 0.413 & 0.725 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 7.66 & 2.45 & 2.1 \\ 0.23 & 1.04 & 0.223 \\ 0.6 & 0.756 & 0.658 \end{bmatrix}$$

$$\text{and} \quad A_0 = \begin{bmatrix} 121 & 18.9 & 15.9 \\ 0 & 2.7 & 0.145 \\ 11.9 & 3.64 & 15.5 \end{bmatrix}.$$

The eigenproblem of $Q(\lambda)$ arose from a study of the oscillations of a wing in an

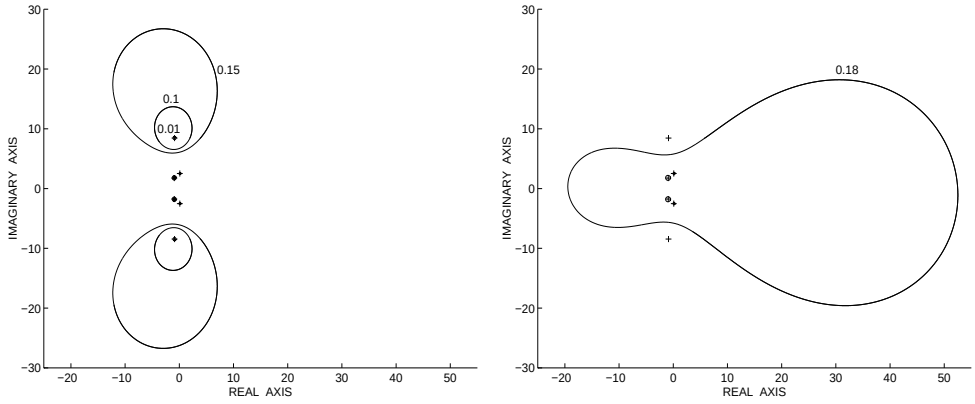

FIG. 5.1. *The boundaries* $\partial\sigma_{\varepsilon,\mathbf{w}}(Q)$ *for* $\varepsilon = 0.01, 0.1, 0.15$ *and for* $\varepsilon = 0.18$.

airstream and originates in [6] (Section 10.11). It has also been examined in Section

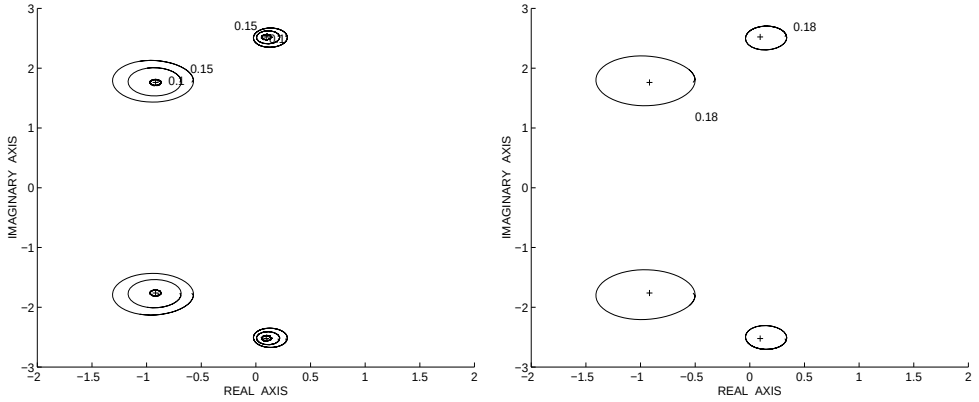

FIG. 5.2. *The less sensitive eigenvalues.*

5.3 of [13], and comparison with the recent treatment of the same example in [22] is interesting. The left part of Figure 5.1 indicates the boundaries of the ε-pseudospectra of $Q(\lambda)$ for $w_0 = w_1 = w_2 = 1$ (i.e., for perturbations measured in the absolute sense) and $\varepsilon = 0.01,\ 0.1,\ 0.15,$ using the present path-tracing algorithm. For step-length $r = 0.1,$ approximately 2000 boundary points are required.

Symmetry of the pseudospectra with respect to the real axis is apparent, confirming Proposition 2.1. The eigenvalues of $Q(\lambda)$ are $-0.88 \pm \mathrm{i}\, 8.44,\ 0.09 \pm \mathrm{i}\, 2.52,\ -0.92 \pm \mathrm{i}\, 1.76,$ and they are plotted as '+'. As observed in [22], the eigenvalues $-0.88 \pm \mathrm{i}\, 8.44$ are seen to be much more sensitive to these perturbations than the remaining eigenvalues of $Q(\lambda)$, and for $\varepsilon < s_{\min}(A_2) \cong 0.17,\ \sigma_{\varepsilon,\mathbf{w}}(Q)$ is bounded and consists of six connected components, confirming Corollary 2.4. On the other hand, in the right part of Figure 5.1, we see that the two connected components corresponding to the pair $-0.88 \pm \mathrm{i}\, 8.44$ become an unbounded connected component with a "hole", where the rest of $\sigma_{\varepsilon,\mathbf{w}}(Q)$ lies, when $\varepsilon = 0.18 > s_{\min}(A_2)$ (i.e., the exterior of the outermost 0.18-curve is in $\sigma_{0.18,\mathbf{w}}(Q)$). The four components of $\sigma_{\varepsilon,\mathbf{w}}(Q)$ that correspond to the eigenvalues $0.09 \pm \mathrm{i}\, 2.52,\ -0.92 \pm \mathrm{i}\, 1.76$ are magnified (using step-length $r = 0.03$) in the left part and the right part of Figure 5.2 for $\varepsilon = 0.01,\ 0.1,\ 0.15$ and for $\varepsilon = 0.18,$ respectively. □

The cost of the algorithm depends mainly on Step 2 since Step 1 usually demands a small number of Newton iterations. In the above example, for initial point $\mu_0 \cong -0.88 + \mathrm{i}\, 8.44,$ search direction $d_0 = 1$ and relative accuracy $tol = 10^{-3},$ the estimation of the first boundary point of $\sigma_{\varepsilon,\mathbf{w}}(Q)$ (for $\varepsilon = 0.1, 0.15, 0.18$) requires only four Newton iterations.

Notice also that in all the examples herein, the matrix polynomials are real and thus we can exploit the symmetry of their pseudospectra with respect to the real axis and need only compute the parts of the boundaries in the closed upper half-plane.

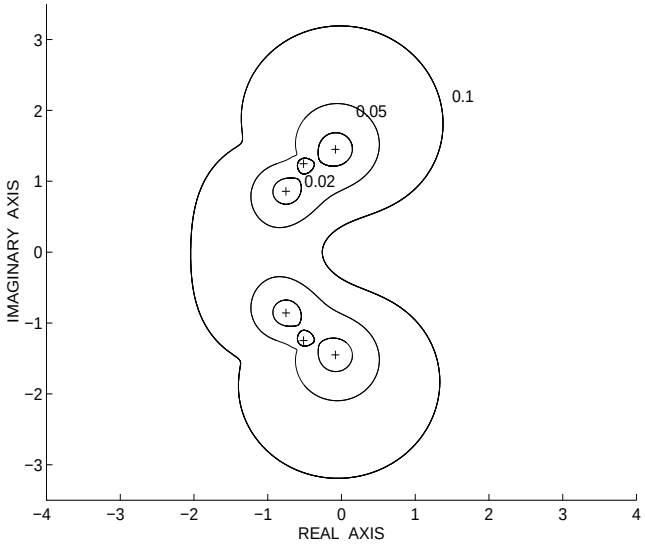

FIG. 5.3. *A damped vibrating system.*

EXAMPLE 5.2. (A vibrating system) The 3×3 selfadjoint matrix polynomial

$$P(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -1 & 6 \end{bmatrix} \lambda + \begin{bmatrix} 2 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 10 \end{bmatrix}$$

corresponds to a mass-spring model described in [5, Example 6]. The predicted boundaries of the ε-pseudospectra of $P(\lambda)$ for $w_0 = \|A_0\| = 10$, $w_1 = \|A_1\| = 6.3$ and $w_2 = \|A_2\| = 5$ (i.e., for perturbations measured in a relative sense) and for $\varepsilon = 0.02, 0.05, 0.1,$ are drawn in Figure 5.3. The eigenvalues of $P(\lambda)$, $-0.08 \pm \mathrm{i}\,1.45$, $-0.75 \pm \mathrm{i}\,0.86$, $-0.51 \pm \mathrm{i}\,1.25,$ are plotted as '+'.

Once again, the predicted symmetry with respect to the real axis is confirmed. Note also that all the pseudospectra are bounded and $s_{\min}(A_2) = 1 > \varepsilon\, w_2,$ verifying Theorem 2.2.

Figure 5.4 shows how the close proximity of two connected components of the pseudospectrum may affect the curve-tracing algorithm. In particular, $\sigma_{0.06,\mathbf{w}}(P)$ consists of two connected components, one in the open upper half-plane and one in the open lower half-plane of $\mathbb{C}$. Using a starting point μ_0 close to the eigenvalue $-0.51 - \mathrm{i}\,1.25,$ direction $d_0 = -1$ and step-length $r = 0.03$, the algorithm sketches a part of the boundary of the lower connected component (in particular, 164 points), and on arriving at the top of the component, it "loses its way" and starts tracing the boundary of the upper connected component, where it remains for ever (see the left part of the figure). If the step-length is decreased to $r = 0.003$, then the algorithm remains on the correct path and plots exactly the boundary of the lower component (right part of the figure). Since $r = 0.003$, approximately 2200 points are needed to

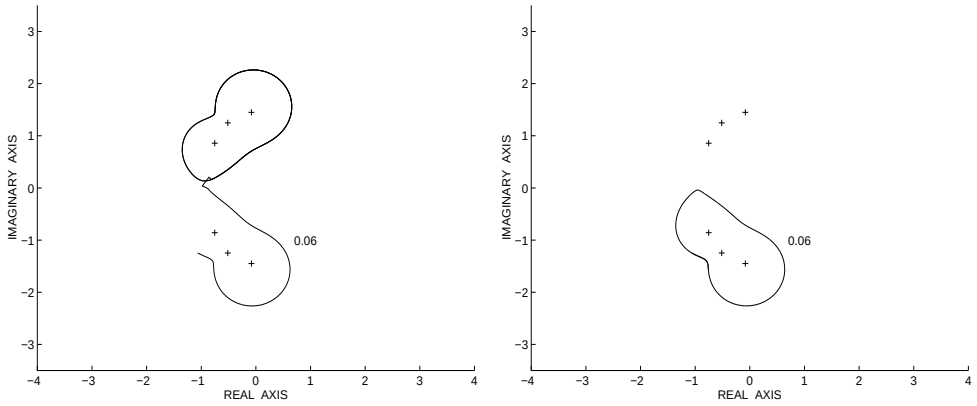

FIG. 5.4. *Step-lengths* $r = 0.03$ *and* $r = 0.003$.

complete the picture. □

In our last example, we consider a quadratic matrix polynomial of larger size.

EXAMPLE 5.3. (A gyroscopic system) Let B denote the 10×10 nilpotent matrix having ones on the subdiagonal and zeros elsewhere, and define the matrices $\hat{M} = (4I_{10} + B + B^T)/6$, $\hat{G} = B - B^T$ and $\hat{K} = B + B^T - 2I_{10}$. Then, using the tensor product, we set

$$\begin{aligned} M &= I_{10} \otimes \hat{M} + 1.30\hat{M} \otimes I_{10}, \\ G &= 1.35I_{10} \otimes \hat{G} + 1.10\hat{G} \otimes I_{10}, \\ K &= I_{10} \otimes \hat{K} + 1.20\hat{K} \otimes I_{10}, \end{aligned}$$

and observe that $M = M^T$, $G = -G^T$ and $K = K^T$. Moreover, the matrix M is positive definite with $s_{\min}(M) \cong 0.8$ and K is negative definite. The 100×100 matrix polynomial $M\lambda^2 + G\lambda + K$ corresponds to an undamped gyroscopic system and its eigenproblem has been examined in [19] (Example 6.1). Adding the tridiagonal damping matrix $D = \text{tridiag}\{-0.1, 0.3, -0.1\}$ to the linear term yields the matrix polynomial

$$R(\lambda) \;=\; M\lambda^2 + (G + D)\lambda + K.$$

The pseudospectra $\sigma_{\varepsilon,\mathbf{w}}(R)$ for $\mathbf{w} = \{1, 1, 1\}$ and for $\varepsilon = 0.004$, 0.02, 0.1, consist of four, two and one bounded connected components, respectively, and their boundaries are drawn in Figure 5.5. The eigenvalues of $R(\lambda)$ are plotted as '+'. As in the two previous examples, the results of Section 2 are apparently confirmed. □

Obviously, the classical grid method of [22] handles several ε's at once, while we have to repeat the path-tracing algorithm for each component and each ε. On the other hand, the new procedure is based on a one-dimensional grid (instead of a predefined two-dimensional grid), and in terms of the number of points at which it is necessary to compute $s_{\min}$, the curve-tracing method competes well with the grid

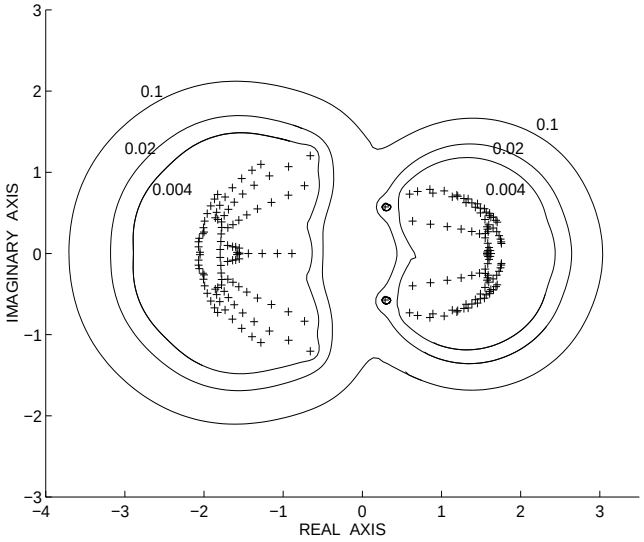

FIG. 5.5. *A damped gyroscopic system.*

method and it is computationally less demanding when seeking a small number of boundary curves. In the above example, for step-length $r = 0.06$, the determination of the boundaries of $\sigma_{0.004,\mathbf{w}}(R)$, $\sigma_{0.02,\mathbf{w}}(R)$ and $\sigma_{0.1,\mathbf{w}}(R)$ in Figure 5.5 requires 292, 304 and 310 points, respectively. For the same step-length, the grid method demands *a priori* knowledge of the size of pseudospectra and a 117×84 grid of the rectangle $\Omega = [-4, 3] \times [-\mathrm{i}\,2.5, \mathrm{i}\,2.5]$, that is, 9828 grid points.

Acknowledgments

The authors are grateful to F. Tisseur and N.J. Higham for helpful comments.

REFERENCES

[1] C. Bekas and E. Gallopoulos, *Cobra: Parallel path following for computing the matrix pseudospectrum*, Parallel Computing, 27 (2001), pp. 1879-1896.

[2] M. Brühl, *A curve tracing algorithm for computing the pseudospectrum*, BIT, 36 (1996), pp. 441-454.

[3] C.-H. Cheng and N. Higham, *The nearest definite pair for the Hermitian generalized eigenvalue problem*, Linear Algebra Appl., 302-303 (1999), pp. 63-76.

[4] M. Embree and L.N. Trefethen, *Generalizing eigenvalue theorems to pseudospectra theorems*, SIAM J. Sci. Comput., 23 (2001), pp. 583-590.

[5] S. Falk, *Klassifikation gedaṁpfter Schwingungssysteme und Eingrenzung ihrer Eigenwerte*, Ingenieur-Archiv, 29 (1960), pp. 436-444.

[6] R.A. Frazer, W.J. Duncan and A.R. Collar, *Elementary Matrices and Some Applications to Dynamics and Differential Equations*, Cambridge University Press, Cambridge, 1938.

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

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

[9] N.J. Higham and F. Tisseur, *More on pseudospectra for polynomial eigenvalue problems and applications in control theory*, Linear Algebra Appl., 351-352 (2002), pp. 435-453.
[10] D. Hinrichsen, N.K. Son and P.H.A. Ngoc, *Stability radii of higher order positive difference systems*, Systems Control Lett., 49 (2003), pp. 377-388.
[11] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
[12] T. Kato, *Perturbation Theory of Linear Operators*, Springer Verlag, Berlin, 1976.
[13] P. Lancaster, *Lambda-Matrices and Vibrating Systems*, Pergamon Press, Oxford, 1966, and Dover Publications, New York, 2002.
[14] C.-K. Li and L. Rodman, *Numerical range of matrix polynomials*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 1256-1265.
[15] S.H. Lui, *Computation of pseudospectra by continuation*, SIAM J. Sci. Comput., 18 (1997), pp. 565-573.
[16] A.S. Markus, *Introduction to the Spectral Theory of Polynomial Operator Pencils*, Amer. Math. Society, Providence, Translations of Math. Monographs, Vol. 71, 1988.
[17] J. Maroulas and P. Psarrakos, *The boundary of numerical range of matrix polynomials*, Linear Algebra Appl., 267 (1997), pp. 101-111.
[18] J. Maroulas and P. Psarrakos, *On the connectedness of numerical range of matrix polynomials*, Linear Algebra Appl., 280 (1998), pp. 97-108.
[19] V. Mehrmann and D. Watkins, *Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltonian/Hamiltonian pencils*, SIAM J. Sci. Comput., 22 (2001), pp. 1905-1925.
[20] R.G. Mosier, *Root neighborhoods of a polynomial*, Math. Comp., 47 (1986), pp. 265-273.
[21] J.-G. Sun, *A note on simple non-zero singular values*, J. Comput. Math., 6 (1988), pp. 258-266.
[22] F. Tisseur and N.J. Higham, *Structured pseudospectra for polynomial eigenvalue problems with applications*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 187-208.
[23] K.-C. Toh and L.N. Trefethen, *Pseudozeros of polynomials and pseudospectra of companion matrices*, Numer. Math., 68 (1994), pp. 403-425.
[24] K.-C. Toh and L.N. Trefethen, *Calculation of pseudospectra by the Arnoldi iteration*, SIAM J. Sci. Comput., 17 (1996), pp. 1-15.
[25] L.N. Trefethen, *Pseudospectra of linear operators*, SIAM Rev., 39 (1997), pp. 383-406.