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, \( \varepsilon \)-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

\[
P(\lambda) = A_m \lambda^m + A_{m-1} \lambda^{m-1} + \cdots + A_1 \lambda + A_0,
\]

where \( \lambda \) 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 \( \lambda_0 \). 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 \( w = \{ w_0, w_1, \ldots, w_m \} \) (written \( w \geq 0 \)), the \( \varepsilon \)-pseudospectrum of \( P(\lambda) \) with respect to \( w \) is defined to be

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

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

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

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

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

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 $w = \{w_0, w_1\} = \{1, 0\}$, then $\sigma_{\varepsilon,w}(P)$ coincides with the $\varepsilon$-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 $\varepsilon$-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 $\varepsilon$-pseudzero 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

$$q_w(\lambda) = w_m \lambda^m + w_{m-1} \lambda^{m-1} + \cdots + w_1 \lambda + w_0.$$

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

$$\sigma_{\varepsilon,w}(P) = \{\lambda \in \mathbb{C} : s_{\min}(P(\lambda)) \leq \varepsilon q_w(|\lambda|)\}.$$  

**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. $\Box$
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 $\varepsilon$-pseudospectrum can be written in the form

$$\partial\sigma_{\varepsilon,w}(P) = \{ \lambda \in \mathbb{C} : s_{\min}(P(\lambda)) = \varepsilon q_w(|\lambda|) \},$$

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

By using efficient algorithms for computing singular values, equations (1.2) and (1.3) become the main tools for estimation of the $\varepsilon$-pseudospectrum of a matrix polynomial. In particular, Tisseur and Higham (see [22]) obtain a graphical representation of $\sigma_{\varepsilon,w}(P)$ 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 $\sigma_{\varepsilon,w}(P)$ are obtained, as well as eigenvalue inclusion properties of the connected components of $\sigma_{\varepsilon,w}(P)$. 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 $w \geq 0$, the $\varepsilon$-pseudospectrum $\sigma_{\varepsilon,w}(P)$ is symmetric with respect to the real axis.

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

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

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

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

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

where $(A_m + \Delta_m)\lambda^m + \cdots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0$ also lies in $B(P, \varepsilon, w)$. Hence, $\mu \in \sigma_{\varepsilon,w}(P)$. 

3
Similarly, if all the coefficients of $P(\lambda)$ are hermitian, then
\[
\det \left[ (A_m + \Delta_m)\lambda^m + \cdots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0 \right]^* =
\]
\[
\det \left[ (A_m + \Delta_m^*)\bar{\lambda}^m + \cdots + (A_1 + \Delta_1^*)\bar{\lambda} + A_0 + \Delta_0^* \right] = 0,
\]
where $(A_m + \Delta_m^*)\lambda^m + \cdots + (A_1 + \Delta_1^*)\lambda + A_0 + \Delta_0^* \in B(P, \varepsilon, w)$. Thus, $\lambda$ lies in $\sigma_{\varepsilon, w}(P)$.

Notice that, if there is a perturbation $P_m(\lambda) \in B(P, \varepsilon, w)$ with identically zero determinant, then $\sigma_{\varepsilon, w}(P)$ coincides with the whole complex plane so, \textit{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 $\varepsilon$ sufficiently small, the pseudospectrum must be bounded and consist of no more than $nm$ connected components. By extending a technique of Li and Rodman \cite[Theorem 2.3]{14}, the next result establishes a necessary and sufficient condition for $\sigma_{\varepsilon, w}(P)$ to be bounded (see also Proposition 3.5 of \cite{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 $\sigma_{\varepsilon, w}(P)$ is bounded if and only if the $\varepsilon w_m$-pseudospectrum of the leading coefficient $A_m$ of $P(\lambda)$ does not contain the origin.

**Proof.** For fixed $\varepsilon > 0$ and $w = \{w_0, w_1, \ldots, w_m\} \geq 0$, 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 $B(P, \varepsilon, w)$ is compact, there is an $M_\varepsilon > 0$ such that for any perturbation of $P(\lambda)$,
\[
P_m(\lambda) = (A_m + \Delta_m)\lambda^m + \cdots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0 \in B(P, \varepsilon, w),
\]
and for any $\lambda \in \mathbb{C}$ with $|\lambda| > M_\varepsilon$, we have
\[
|\det P_m(\lambda) - \det(A_m + \Delta_m)\lambda^m| < \zeta_\varepsilon|\lambda|^m \leq |\det(A_m + \Delta_m)\lambda^m|,
\]
(keeping in mind that $\det P_m(\lambda)$ is a scalar polynomial with leading term $\det(A_m + \Delta_m)\lambda^m$). Hence, $\det P_m(\lambda) \neq 0$, i.e., $\sigma_{\varepsilon, w}(P) \subseteq \{\lambda \in \mathbb{C} : |\lambda| \leq M_\varepsilon\}; \sigma_{\varepsilon, w}(P)$ is bounded.

To prove the converse, assume that $\sigma_{\varepsilon, w}(P)$ is bounded but there is a $P_m(\lambda) = (A_m + \Delta_m)\lambda^m + \cdots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0 \in B(P, \varepsilon, w)$ with $\det(A_m + \Delta_m) = 0$. Then at least one of the coefficients of the scalar polynomial $\det P_m(\lambda)$ is nonzero: otherwise $\sigma_{\varepsilon, w}(P) = \mathbb{C}$, a contradiction. Let the coefficient of $\lambda^\tau$ ($\tau \in \{0, 1, \ldots, nm - 1\}$) in $\det P_m(\lambda)$ be nonzero and denote it by $\beta_\tau$. Construct a sequence $\{\Delta_{m,k}\}_{k \in \mathbb{N}} \subset \mathbb{C}^{n \times n}$ such that $\lim_{k \to \infty} \Delta_{m,k} = \Delta_m$ and
\[
\det(A_m + \Delta_{m,k}) \neq 0 \quad \text{and} \quad \|\Delta_{m,k}\| \leq \varepsilon w_m \quad (k = 1, 2, \ldots).
\]
Clearly, for a fixed $\delta > 0$, $|\det(A_m + \Delta_{m,k})| < \delta$ for all sufficiently large $k$. Since $\sigma_{e,w}(P)$ is bounded, the $(nm - \tau)$th elementary symmetric function of the roots of $\text{det}[(A_m + \Delta_{m,k})\lambda^m + \cdots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0]$, which is equal to $\pm \beta / \det(A_m + \Delta_{m,k})$, is bounded for all $k$. This contradicts the construction of the sequence $\{\Delta_{m,k}\}_{k \in \mathbb{N}}$. □

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

**Theorem 2.3.** If $\sigma_{e,w}(P)$ is bounded, then it has no more than $nm$ connected components, and any $P_\Delta(\lambda) \in \mathcal{B}(P, \varepsilon, w)$ 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 $\sigma_{e,w}(P)$.

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

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

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 $\mathcal{B}(P, \varepsilon, w)$ and its eigenvalues lie in $\sigma_{e,w}(P)$.

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 $\sigma_{e,w}(P)$ and its $nm - k$ remaining eigenvalues are isolated in $\sigma_{e,w}(P) \setminus \mathcal{G}$, 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 $\sigma_{e,w}(P)$ 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, $\sigma_{e,w}(P)$ cannot have more than $nm$ bounded connected components. □

**Corollary 2.4.** For any $\varepsilon > 0$ such that $\varepsilon w_m < s_{\min}(A_m)$, $\sigma_{e,w}(P)$ 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 $\varepsilon$-pseudospectrum of $A$ lies in the region

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

where $\text{dist}(\lambda, F(A))$ denotes the distance between the point $\lambda$ 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 $nc(\mathcal{G})$ 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 $\xi$ ($\leq m$) connected components $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_\xi$. Then for given $\varepsilon > 0$ and $w \geq 0$,

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

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

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

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

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

$$v_0^*\Delta_{\lambda_0}^m + \cdots + \Delta_1\lambda_0 + \Delta_0)v_0 = -v_0^*P(\lambda_0)v_0,$$

and consequently,

$$\sum_{j=0}^m |v_0^*\Delta_jv_0|^2 |\lambda_0|^{2j} \geq \sum_{j=0}^m (v_0^*\Delta_jv_0) \lambda_0^{j} = |v_0^*P(\lambda_0)v_0|$$

$$= |v_0^*A_mv_0| \prod_{j=1}^m |\lambda_0 - \lambda_j(v_0)|,$$

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 $\varepsilon w_j \geq \|\Delta_j\| > |v_0^*\Delta_jv_0|$, $j = 0, 1, \ldots, m$, it follows that

$$\varepsilon q_w(|\lambda_0|) \geq \sum_{j=0}^m \|\Delta_j\| |\lambda_0|^{2j} \geq \sum_{j=0}^m |v_0^*\Delta_jv_0| |\lambda_0|^j.$$
\[ \geq \tilde{r}(A_m) \prod_{j=1}^{m} |\lambda_0 - \lambda_j(v_0)| \geq \tilde{r}(A_m) \prod_{j=1}^{\xi} \text{dist}(\lambda_0, G_j)^{s(G_j)}. \]

The proof is complete. \[\Box\]

4. A curve-tracing algorithm. For fixed \( \varepsilon > 0 \) and \( w = \{w_0, w_1, \ldots, w_m\} \geq 0 \), the boundary of the \( \varepsilon \)-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 \( \partial \sigma_{\varepsilon, w}(P) \) 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

\[ g_P(x, y) = s_{\min}(P(x + iy)) : x, y \in \mathbb{R}. \]

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

Then there is a neighborhood \( \mathcal{N} \) of \( b_0 \) on which a simple nonzero singular value \( s(b) \) of \( F(b) \) is defined together with corresponding left and right singular vectors \( u(b) \) and \( v(b) \), respectively, such that \( s(b_0) = s_0 \), \( u(b_0) = u_0 \) and \( v(b_0) = v_0 \), and the functions \( s, u \) and \( v \) are real analytic on \( \mathcal{N} \).

Moreover, the partial derivatives of \( s(b) \) at \( b_0 \) are given by

\[ \frac{\partial s(b_0)}{\partial b^{(j)}} = \text{Re} \left( u_0^* \frac{\partial F(b_0)}{\partial b^{(j)}} v_0 \right) ; \quad j = 1, 2, \ldots, d. \]

The next corollary is a direct consequence of this theorem.

**Corollary 4.2.** Let \( \lambda_0 = x_0 + i y_0 \in \mathbb{C} \setminus \sigma(P) \). If \( s_{\min}(P(\lambda_0)) \) is a simple singular value of \( P(\lambda_0) \), 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

\[ \nabla g_P(x_0, y_0) = \left( \text{Re} \left( u_0^* \frac{\partial P(x_0 + iy_0)}{\partial x} v_0 \right), \text{Re} \left( u_0^* \frac{\partial P(x_0 + iy_0)}{\partial y} v_0 \right) \right). \]

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

\[ \partial \sigma_{\varepsilon, w}(P) = \{ \lambda \in \mathbb{C} : g_P(\lambda) - \varepsilon q_w(|\lambda|) = 0 \}, \]
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 $\sigma(P)$, let $\lambda_0 = x_0 + iy_0 \in \sigma(P) \setminus \sigma$ and $d_0 \in \mathbb{C}$ be nonzero, and consider the function

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

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 $\lambda_0$ and the gradient $\nabla g_P(x_0, y_0)$ given by Corollary 4.2 is nonzero.

Then the first Newton iterate is

$$t_1 = t_0 - \frac{h(t_0)}{h'(t_0)} = -\frac{g_P(\lambda_0) - \varepsilon q_w(|\lambda_0|)}{h'(0)};$$

where

$$h'(0) = \frac{\partial [g_P(\lambda_0) - \varepsilon q_w(|\lambda_0|)]}{\partial d_0} = \frac{\partial [g_P(x_0 + iy_0) - \varepsilon q_w(\sqrt{x_0^2 + y_0^2})]}{\partial d_0}$$

$$= (\text{Re } d_0) \left[ \text{Re} \left( u^*_0 \frac{\partial P(x_0 + iy_0)}{\partial x} v_0 \right) - \varepsilon \frac{\partial q_w(\sqrt{x_0^2 + y_0^2})}{\partial x} \right]$$

$$+ (\text{Im } d_0) \left[ \text{Re} \left( u^*_0 \frac{\partial P(x_0 + iy_0)}{\partial y} v_0 \right) - \varepsilon \frac{\partial q_w(\sqrt{x_0^2 + y_0^2})}{\partial y} \right].$$

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

$$\text{Re} \left( u^*_0 \frac{\partial P(x_0 + iy_0)}{\partial x} v_0 \right) - \varepsilon \frac{\partial q_w(\sqrt{x_0^2 + y_0^2})}{\partial x} = R_{\varepsilon,w}(\lambda_0, x)$$

and

$$\text{Re} \left( u^*_0 \frac{\partial P(x_0 + iy_0)}{\partial y} v_0 \right) - \varepsilon \frac{\partial q_w(\sqrt{x_0^2 + y_0^2})}{\partial y} = R_{\varepsilon,w}(\lambda_0, y).$$

Then

$$(4.1) z_1 = \lambda_0 + t_1 d_0 = x_0 + iy_0 - \frac{g_P(x_0 + iy_0) - \varepsilon q_w(\sqrt{x_0^2 + y_0^2})}{(\text{Re } d_0) R_{\varepsilon,w}(\lambda_0, x) + (\text{Im } d_0) R_{\varepsilon,w}(\lambda_0, y)} d_0.$$

Since the point $\lambda_0 = x_0 + iy_0$ lies in $\sigma(P)$, for suitably chosen direction $d_0$, we can estimate a boundary point of $\sigma(P)$ (in particular, an intersection point of $\partial \sigma(P)$
and \( \{ \lambda_0 + t d_0 : t \in \mathbb{R} \} \) by repeating (4.1) until the quantity \( |s_{\min}(P(z)) - \varepsilon q_w(|z|)| \) is small enough.

It is also worth noting that

\[
(\text{Re} \, d_0) R_{\varepsilon, w}(\lambda_0, x) + (\text{Im} \, d_0) R_{\varepsilon, w}(\lambda_0, y) = \text{Re} \left[ d_0 (R_{\varepsilon, w}(\lambda_0, x) + iR_{\varepsilon, w}(\lambda_0, y)) \right],
\]

and on choosing the direction \( d_0 = R_{\varepsilon, w}(\lambda_0, x) + iR_{\varepsilon, w}(\lambda_0, y) \), the equation (4.1) implies

\[
(4.2) \quad z_1 = x_0 + iy_0 - \frac{g_P(x_0 + iy_0) - \varepsilon q_w \left( \sqrt{x_0^2 + y_0^2} \right)}{R_{\varepsilon, w}(\lambda_0, x) - iR_{\varepsilon, w}(\lambda_0, y)}.
\]

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

\[
\hat{z}_k = z_{k-1} + \tau_k \left( i \frac{\nabla g_P(z_{k-1}) - \varepsilon q_w(|z_{k-1}|)}{\left| \nabla g_P(z_{k-1}) - \varepsilon q_w(|z_{k-1}|) \right|} \right) = z_{k-1} + \tau_k \left( i \frac{R_{\varepsilon, w}(z_{k-1}, x), R_{\varepsilon, w}(z_{k-1}, y)}{|R_{\varepsilon, w}(z_{k-1}, x), R_{\varepsilon, w}(z_{k-1}, y)|} \right),
\]

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

**Correction:** For sufficiently small \( \tau_k \), the correction step consists of a single Newton iterate with respect to the equation \( g_P(z_k + t d_k) - \varepsilon q_w(|z_k + t d_k|) = 0 \), 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

\[
\nabla [g_P(z_k) - \varepsilon q_w(|z_k|)] = (R_{\varepsilon, w}(\hat{z}_k, x), R_{\varepsilon, w}(\hat{z}_k, y)).
\]

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

\[
\nabla [g_P(z_{k-1}) - \varepsilon q_w(|z_{k-1}|)] = (R_{\varepsilon, w}(z_{k-1}, x), R_{\varepsilon, w}(z_{k-1}, y))
\]

and (4.2) is written in the form

\[
(4.4) \quad z_k = \hat{z}_k - \frac{g_P(z_{k-1}) - \varepsilon q_w(|z_{k-1}|)}{R_{\varepsilon, w}(z_{k-1}, x) - iR_{\varepsilon, w}(z_{k-1}, y)}.
\]

In Figure 4.1, a graphical illustration of the directions \( d_k = \nabla [g_P(z_k) - \varepsilon q_w(|z_k|)] \) (left part) and \( d_k = \nabla [g_P(z_{k-1}) - \varepsilon q_w(|z_{k-1}|)] \) (right part) is given. The prediction points are marked with asterisks, the correction points are plotted as ‘o’, and \( \nabla \) denotes the gradient \( \nabla [g_P(z_{k-1}) - \varepsilon q_w(|z_{k-1}|)] \). Note that in the right part of the figure, the line segment \([\hat{z}_k, z_k]\) is parallel to \( \nabla \).
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 $\partial \sigma_{\varepsilon, w}(P)$, an approximation $\mu_0$ 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 $|s_{\min}(P(z_{\text{new}})) - \varepsilon q_w(|z_{\text{new}}|)| > \varepsilon q_w(|z_{\text{new}}|) tol$, 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

$$\nabla [g_P(z_{\text{old}}) - \varepsilon q_w(|z_{\text{old}}|)] = (R_{\varepsilon, w}(z_{\text{old}}, x), R_{\varepsilon, w}(z_{\text{old}}, y)).$$

(d) Compute the next Newton iterate

$$z_{\text{new}} = z_{\text{old}} - \frac{s_{\min}(P(z_{\text{old}})) - \varepsilon q_w(|z_{\text{old}}|)}{(Re d_0) R_{\varepsilon, w}(z_{\text{old}}, x) + (Im d_0) R_{\varepsilon, w}(z_{\text{old}}, y)} d_0.$$

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

$$\nabla \left[\frac{g_P(z_{k-1}) - \varepsilon q_w(|z_{k-1}|)}{R_{\varepsilon, w}(z_{k-1}, x) - i R_{\varepsilon, w}(z_{k-1}, y)}\right]$$

and the direction

$$d_k = \frac{i R_{\varepsilon, w}(z_{k-1}, x) + i R_{\varepsilon, w}(z_{k-1}, y)}{|R_{\varepsilon, w}(z_{k-1}, x) + i R_{\varepsilon, w}(z_{k-1}, y)|}.$$

(c) Compute the predicted point

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

(d) Compute the corrected point

$$z_k = \hat{z}_k - \frac{s_{\min}(P(z_{k-1})) - \varepsilon q_w(|z_{k-1}|)}{R_{\varepsilon, w}(z_{k-1}, x) - i R_{\varepsilon, w}(z_{k-1}, y)}.$$

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 $s_{\varepsilon, w}(P)$ containing $\mu_0$. Consequently, for a complete picture of $s_{\varepsilon, w}(P)$, it may be necessary to repeat the procedure several times with different (appropriate) values of $\mu_0$.

Since the size of the step-lengths $\tau_k (k = 2, 3, \ldots, N)$ 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 $\tau_k$ will undoubtedly depend mainly on the local shape of the curve $\partial s_{\varepsilon, w}(P)$. 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 $\partial s_{\varepsilon, w}(P)$ 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 $\partial s_{\varepsilon, w}(P)$ 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 $\mu_0$ 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 $\sigma_{\varepsilon,w}(P)$, 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

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

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

The eigenproblem of $Q(\lambda)$ arose from a study of the oscillations of a wing in an airstream and originates in [6] (Section 10.11). It has also been examined in Section

![Fig. 5.1. The boundaries $\partial\sigma_{\varepsilon,w}(Q)$ for $\varepsilon = 0.01, 0.1, 0.15$ and for $\varepsilon = 0.18$.](image)
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 \( \varepsilon \)-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 \(-0.88 \pm i 8.44\), \(0.09 \pm i 2.52\), \(-0.92 \pm i 1.76\), and they are plotted as ‘+’. As observed in [22], the eigenvalues \(-0.88 \pm i 8.44\) are seen to be much more sensitive to these perturbations than the remaining eigenvalues of \( Q(\lambda) \), and for \( \varepsilon < s_{\text{min}}(A_2) \approx 0.17 \), \( \sigma_{\varepsilon,w}(Q) \) 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 \(-0.88 \pm i 8.44\) become an unbounded connected component with a “hole”, where the rest of \( \sigma_{\varepsilon,w}(Q) \) lies, when \( \varepsilon = 0.18 > s_{\text{min}}(A_2) \) (i.e., the exterior of the outermost 0.18-curve is in \( \sigma_{0.18,w}(Q) \)). The four components of \( \sigma_{\varepsilon,w}(Q) \) that correspond to the eigenvalues \(0.09 \pm i 2.52\), \(-0.92 \pm i 1.76\) 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 \( \mu_0 = -0.88 + i 8.44 \), search direction \( d_0 = 1 \) and relative accuracy \( tol = 10^{-3} \), the estimation of the first boundary point of \( \sigma_{\varepsilon,w}(Q) \) (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.
Example 5.2. (A vibrating system) The $3 \times 3$ 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 $\varepsilon$-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 $P(\lambda)$, $-0.08 \pm 1.145$, $-0.75 \pm 10.86$, $-0.51 \pm 1.25$, 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, $\sigma_{0.06,w}(P)$ 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 $\mu_0$ close to the eigenvalue $-0.51 - 1.25$, 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
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 $10 \times 10$ 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

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

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 $100 \times 100$ 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 $\sigma_{\varepsilon,w}(R)$ for $w = \{1, 1, 1\}$ 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 $\varepsilon$’s at once, while we have to repeat the path-tracing algorithm for each component and each $\varepsilon$. 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
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 $\sigma_{0.004,w}(R)$, $\sigma_{0.02,w}(R)$ and $\sigma_{0.1,w}(R)$ 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 $117 \times 84$ grid of the rectangle $\Omega = [-4,3] \times [-12.5,2.5]$, that is, 9828 grid points.

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

REFERENCES