Structured pseudospectra for polynomial eigenvalue problems, with applications

Tisseur, Françoise and Higham, Nicholas J.

2001

MIMS EPrint: **2006.144**

Manchester Institute for Mathematical Sciences

School of Mathematics

The University of Manchester

STRUCTURED PSEUDOSPECTRA FOR POLYNOMIAL EIGENVALUE PROBLEMS, WITH APPLICATIONS*

FRANÇOISE TISSEUR† AND NICHOLAS J. HIGHAM†

Abstract. Pseudospectra associated with the standard and generalized eigenvalue problems have been widely investigated in recent years. We extend the usual definitions in two respects, by treating the polynomial eigenvalue problem and by allowing structured perturbations of a type arising in control theory. We explore connections between structured pseudospectra, structured backward errors, and structured stability radii. Two main approaches for computing pseudospectra are described. One is based on a transfer function and employs a generalized Schur decomposition of the companion form pencil. The other, specific to quadratic polynomials, finds a solvent of the associated quadratic matrix equation and thereby factorizes the quadratic λ-matrix. Possible approaches for large, sparse problems are also outlined. A collection of examples from vibrating systems, control theory, acoustics, and fluid mechanics is given to illustrate the techniques.

Key words. polynomial eigenvalue problem, λ-matrix, matrix polynomial, pseudospectrum, stability radius, backward error, transfer function, quadratic matrix equation, solvent, structured perturbations, Orr–Sommerfeld equation

AMS subject classifications. 65F15, 15A22

PII. S0895479800371451

1. Introduction. Pseudospectra are an established tool for gaining insight into the sensitivity of the eigenvalues of a matrix to perturbations. Their use is widespread with applications in areas such as fluid mechanics, Markov chains, and control theory. Most of the existing work is for the standard eigenproblem, although attention has also been given to matrix pencils [4], [23], [33], [40], [46]. The literature on pseudospectra is large and growing. We refer to Trefethen [41], [42], [43] for thorough surveys of pseudospectra and their computation for a single matrix; see also the Web site [3].

In this work we investigate pseudospectra for polynomial matrices (or λ-matrices)

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

where $A_k \in \mathbb{C}^{n \times n}$, $k = 0 : m$. We first define the ϵ-pseudospectrum and obtain a computationally useful characterization. We examine the relation between the backward error of an approximate eigenpair of the polynomial eigenvalue problem associated with (1.1), the ϵ-pseudospectrum, and the stability radius. We consider both unstructured perturbations and structured perturbations of a type commonly used in control theory.

Existing methods for the computation of pseudospectra in the case $m = 1$ (the standard and generalized eigenvalue problems) do not generalize straightforwardly to matrix polynomials. We develop two techniques that allow efficient computation for $m > 1$. A transfer function approach employs the generalized Schur decomposition of the $mn \times mn$ companion form pencil. For the quadratic case ($m = 2$) an alternative

solvent approach computes a solvent of the associated quadratic matrix equation $A_2 X^2 + A_1 X + A_0 = 0$ and thereby factorizes the quadratic λ-matrix; it works all the time with $n \times n$ matrices once the solvent has been obtained. We give a detailed comparison of these approaches and also outline techniques that can be efficiently used when n is so large as to preclude factorizations.

In the last section, we illustrate our theory and techniques on applications from vibrating systems, control theory, acoustics, and fluid mechanics.

2. Pseudospectra.

2.1. Definition. The polynomial eigenvalue problem is to find the solutions (x, λ) of

$$(2.1) \qquad P(\lambda)x = 0,$$

where $P(\lambda)$ is of the form (1.1). If $x \neq 0$ then λ is called an eigenvalue and x the corresponding right eigenvector; $y \neq 0$ is a left eigenvector if $y^* P(\lambda) = 0$. The set of eigenvalues of P is denoted by $\Lambda(P)$. When A_m is nonsingular P has mn finite eigenvalues, while if A_m is singular P has infinite eigenvalues. Good references for the theory of λ-matrices are [8], [20], [21], [37].

Throughout this paper we assume that P has only finite eigenvalues (and pseudoeigenvalues); how to deal with infinite eigenvalues is described in [16].

For notational convenience, we introduce

$$(2.2) \qquad \Delta P(\lambda) = \lambda^m \Delta A_m + \lambda^{m-1} \Delta A_{m-1} + \cdots + \Delta A_0.$$

We define the ϵ-pseudospectrum of P by

$$\Lambda_\epsilon(P) = \big\{ \lambda \in \mathbb{C} : (P(\lambda) + \Delta P(\lambda))x = 0 \text{ for some } x \neq 0 \text{ and } \Delta P(\lambda)$$
$$(2.3) \qquad\qquad \text{with } \|\Delta A_k\| \leq \epsilon \alpha_k, \ k = 0{:}m \big\}.$$

Here the α_k are nonnegative parameters that allow freedom in how perturbations are measured—for example, in an absolute sense ($\alpha_k \equiv 1$) or a relative sense ($\alpha_k = \|A_k\|$). By setting $\alpha_k = 0$ we can force $\Delta A_k = 0$ and thus keep A_k unperturbed. The norm, here and throughout, is any subordinate matrix norm. Occasionally, we will specialize to the norm $\|\cdot\|_p$ subordinate to the Hölder vector p-norm.

When $P(\lambda) = A - \lambda I$, $\Delta P(\lambda) = \Delta A$ and $\alpha_1 = 1$, definition (2.3) reduces to the standard definition of ϵ-pseudospectrum of a single matrix:

$$(2.4) \qquad \Lambda_\epsilon(A) = \big\{ \lambda \in \mathbb{C} : \lambda \in \Lambda(A + \Delta A) \text{ for some } \Delta A \text{ with } \|\Delta A\| \leq \epsilon \big\}.$$

It is well known [43] that (2.4) is equivalent to

$$\Lambda_\epsilon(A) = \big\{ \lambda \in \mathbb{C} : \|(\lambda I - A)^{-1}\| \geq \epsilon^{-1} \big\}.$$

In the following lemma, we provide a generalization of this equivalence for the ϵ-pseudospectrum of P.

LEMMA 2.1.

$$\Lambda_\epsilon(P) = \big\{ \lambda \in \mathbb{C} : \|P(\lambda)^{-1}\| \geq (\epsilon\, p(|\lambda|))^{-1} \big\},$$

where $p(x) = \sum_{k=0}^m \alpha_k x^k$.

Proof. Let $\mathcal{S}$ denote the set on the right-hand side of the claimed equality. We first show that $\lambda \in \Lambda_\epsilon(P)$ implies $\lambda \in \mathcal{S}$. If λ is an eigenvalue of P this is immediate, so we can assume that λ is not an eigenvalue of P and hence that $P(\lambda)$ is nonsingular. Since

$$P(\lambda) + \Delta P(\lambda) = P(\lambda)\big(I + P(\lambda)^{-1}\Delta P(\lambda)\big)$$

is singular, we have

$$1 \le \|P(\lambda)^{-1}\Delta P(\lambda)\| \le \|P(\lambda)^{-1}\| \left(\sum_{k=0}^m |\lambda|^k \alpha_k \epsilon\right) = \|P(\lambda)^{-1}\|\, \epsilon\, p(|\lambda|),$$

so that $\lambda \in \mathcal{S}$.

Now let $\lambda \in \mathcal{S}$. Again we can assume that $P = P(\lambda)$ is nonsingular. Choose y with $\|y\| = 1$ so that $\|P^{-1}y\| = \|P^{-1}\|$ and let $x = P^{-1}y/\|P^{-1}\|$, so that $\|x\| = 1$. Then there exists a matrix H with $\|H\| = 1$ such that $Hx = y$ (see, for example, [11, Lem. 6.3]). Let $E = -H/\|P^{-1}\|$. Then

$$(P + E)x = \frac{y}{\|P^{-1}\|} - \frac{y}{\|P^{-1}\|} = 0$$

and

$$\|E\| = 1/\|P^{-1}\| \le \epsilon\, p(|\lambda|).$$

We now apportion E between the A_k by defining

$$\Delta A_k = \mathrm{sign}(\lambda^k)\alpha_k p(|\lambda|)^{-1}E,$$

where for complex z we define

$$\mathrm{sign}(z) = \begin{cases} \overline{z}/|z|, & z \ne 0, \\ 0, & z = 0. \end{cases}$$

Then

$$\Delta P(\lambda) = \sum_{k=0}^m \lambda^k \Delta A_k = \left(\sum_{k=0}^m |\lambda|^k \alpha_k\right) p(|\lambda|)^{-1}E = E$$

and $\|\Delta A_k\| \le \alpha_k \epsilon$, $k = 0{:}\, m$. Hence $\lambda \in \Lambda_\epsilon(P)$. $\square$

The characterization of the ϵ-pseudospectrum in Lemma 2.1 will be the basis of our algorithms for computing pseudospectra.

We note that for $n = 1$, $\Lambda_\epsilon(P)$ is the root neighborhood of the polynomial P introduced by Mosier [28], that is, the set of all polynomials obtained by elementwise perturbations of P of size at most ϵ. This set is also investigated by Toh and Trefethen [38], who call it the ϵ-pseudozero set.

2.2. Connection with backward error. A natural definition of the normwise backward error of an approximate eigenpair (x, λ) of (2.1) is

(2.5) $\qquad \eta(x, \lambda) := \min\{\, \epsilon : (P(\lambda) + \Delta P(\lambda))x = 0, \ \|\Delta A_k\| \le \epsilon\alpha_k, \ k = 0{:}\, m \,\},$

and the backward error for an approximate eigenvalue λ is given by

(2.6) $\qquad\qquad\qquad\qquad \eta(\lambda) := \min_{x \ne 0} \eta(x, \lambda).$

By comparing the definitions (2.3) and (2.6) it is clear that the ϵ-pseudospectrum can be expressed in terms of the backward error of λ as

$$(2.7) \qquad \Lambda_\epsilon(P) = \{\, \lambda \in \mathbb{C} : \eta(\lambda) \leq \epsilon \,\}.$$

The following lemma gives an explicit expression for $\eta(x, \lambda)$ and $\eta(\lambda)$. This lemma generalizes results given in [36] for the 2-norm and earlier in [5], [10] for the generalized eigenvalue problem.

LEMMA 2.2. *The normwise backward error $\eta(x, \lambda)$ is given for $x \neq 0$ by*

$$(2.8) \qquad \eta(x, \lambda) = \frac{\|r\|}{p(|\lambda|)\|x\|},$$

where $r = P(\lambda)x$ and $p(x) = \sum_{k=0}^{m} \alpha_k x^k$. If λ is not an eigenvalue of P then

$$(2.9) \qquad \eta(\lambda) = \frac{1}{p(|\lambda|)\|P(\lambda)^{-1}\|}.$$

Proof. It is straightforward to show that the right-hand side of (2.8) is a lower bound for $\eta(x, \lambda)$. That the lower bound is attained is proved using a construction for ΔA_k similar to that in the proof of Lemma 2.1. The expression (2.9) follows on using the equality, for nonsingular $C \in \mathbb{C}^{n \times n}$, $\min_{x \neq 0} \|Cx\|/\|x\| = \|C^{-1}\|^{-1}$. ☐

We observe that the expressions (2.7) and (2.9) lead to another proof of Lemma 2.1.

2.3. Structured perturbations. We now suppose that $P(\lambda)$ is subject to structured perturbations that can be expressed as

$$(2.10) \qquad [\, \Delta A_0, \quad \ldots, \quad \Delta A_m \,] = D\Theta\,[\, E_0, \quad \ldots, \quad E_m \,],$$

with $D \in \mathbb{C}^{n \times s}$, $\Theta \in \mathbb{C}^{s \times t}$, and $E = [\, E_0, \quad \ldots, \quad E_m \,] \in \mathbb{C}^{t \times n(m+1)}$. The matrices D and E are fixed and assumed to be of full rank, and they define the structure of the perturbations; Θ is an arbitrary matrix whose elements are the free parameters. Note that $\Delta A_0, \ldots, \Delta A_m$ in (2.10) are linear functions of the parameters in Θ, but that not all linear functions can be represented in this form. We choose this particular structure for the perturbations because it is one commonly used in control theory [17], [18], [30] and it leads to more tractable formulae than a fully general approach. Note, for instance, that the system

$$\dot{x}(t) = (A + D\Theta E)x(t), \quad t > 0$$

(which leads to a polynomial eigenvalue problem with $m = 1$), may be interpreted as a closed loop system with unknown static linear output feedback Θ; see Figure 2.1.

Note that unstructured perturbations are represented by the special case of (2.10) with

$$(2.11) \quad s = n, \quad t = n(m+1), \quad D = I_n, \quad \Theta = [\, \Delta A_0, \ldots, \Delta A_m \,], \quad E = I_{n(m+1)}.$$

For notational convenience, we introduce

$$E(\lambda) = E\,[\, I_n, \quad \lambda I_n, \quad \ldots, \quad \lambda^m I_n \,]^T = \lambda^m E_m + \lambda^{m-1} E_{m-1} + \cdots + E_0.$$

Corresponding to (2.10) we have the following definition of structured backward error for an approximate eigenpair (x, λ):

$$\eta(x, \lambda; D, E) := \min_{\Theta \in \mathbb{C}^{s \times t}} \{\, \|\Theta\| : (P(\lambda) + \Delta P(\lambda))x = 0, \; \Delta P(\lambda) = D\Theta E(\lambda) \,\},$$
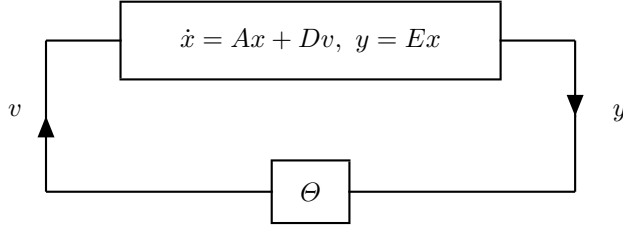
$$\dot{x} = Ax + Dv, \ y = Ex$$

$$v \qquad\qquad\qquad y$$

$$\Theta$$

FIG. 2.1. *Closed loop system with unknown static linear output feedback Θ.*

and the backward error for an approximate eigenvalue is

$$\eta(\lambda; D, E) := \min_{x \neq 0} \eta(x, \lambda; D, E).$$

In the next result we use a superscript "+" to denote the pseudo-inverse [9].

LEMMA 2.3. *The structured backward error $\eta(x, \lambda; D, E)$ in the Frobenius norm is given by*

$$(2.12) \qquad\qquad \eta_F(x, \lambda; D, E) = \|D^+ P(\lambda) \, x \, [E(\lambda)x]^+\|_F$$

if the system

$$(2.13) \qquad\qquad D\Theta E(\lambda)x = -P(\lambda)x$$

is consistent; otherwise $\eta_F(x, \lambda; D, E)$ is infinite.

Proof. It is immediate that $\eta_F(x, \lambda; D, E)$ is the Frobenius norm of the minimum Frobenius norm solution to (2.13). The result follows from the fact that $X = A^+CB^+$ is the solution of minimum Frobenius norm to the consistent system $AXB = C$ [31, sect. 3.4.8]. □

To gain some insight into the expression (2.12) we consider the case of unstructured but weighted perturbations, as in (2.11) but with

$$E = \text{diag}(\alpha_0 I_n, \ldots, \alpha_m I_n) =: I_{n(m+1)}^\alpha, \qquad E(\lambda) = [\, \alpha_0 I_n, \quad \ldots, \quad \alpha_m \lambda^m I_n \,]^T .$$

The system (2.13) is now trivially consistent and (2.12) gives

$$\eta_F(x, \lambda; I_n, I_{n(m+1)}^\alpha) = \left\| P(\lambda)x \begin{bmatrix} \alpha_0 x \\ \vdots \\ \alpha_m \lambda^m x \end{bmatrix}^+ \right\|_F$$

$$(2.14) \qquad\qquad = \frac{\|P(\lambda)x\|_2}{\left(\sum_{i=0}^m \alpha_i^2 |\lambda|^{2i}\right)^{1/2} \|x\|_2},$$

using the fact that $\|ab^*\|_F = \|a\|_2\|b\|_2$ for $a, b \in \mathbb{C}^n$. The expression (2.14) differs from that for $\eta(x, \lambda)$ in (2.8) for the 2-norm only by having the 2-norm of the vector $[\alpha_0 \ \ldots \ \alpha_m \lambda^m]$ rather than the 1-norm in the denominator.

LEMMA 2.4. *If λ is not an eigenvalue of $P(\lambda)$ then the structured backward error $\eta(\lambda; D, E)$ is given by*

$$(2.15) \qquad\qquad \eta(\lambda; D, E) = \|E(\lambda)P(\lambda)^{-1}D\|^{-1}.$$

Proof. We have

$$
\begin{aligned}
\eta(\lambda; D, E) &= \min_{x \neq 0} \eta(x, \lambda; D, E) \\
&= \min_{x \neq 0} \min_{\Theta \in \mathbb{C}^{s \times t}} \big\{ \|\Theta\| : (P(\lambda) + \Delta P(\lambda))x = 0,\ \Delta P(\lambda) = D\Theta E(\lambda) \big\} \\
&= \min_{\Theta \in \mathbb{C}^{s \times t}} \big\{ \|\Theta\| : \det(P(\lambda) + \Delta P(\lambda)) = 0,\ \Delta P(\lambda) = D\Theta E(\lambda) \big\}.
\end{aligned}
$$

The companion form of $P(\lambda) + \Delta P(\lambda)$ is given by

$$
F - \lambda G + \Delta F - \lambda \Delta G,
$$

where

$$
(2.16)\quad F = \begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \ddots & \vdots \\ \vdots & & & \ddots & 0 \\ & & & & I \\ -A_0 & -A_1 & -A_2 & \cdots & -A_{m-1} \end{bmatrix}, \qquad G = \begin{bmatrix} I & & & & \\ & I & & & \\ & & \ddots & & \\ & & & I & \\ & & & & A_m \end{bmatrix},
$$

and

$$
\Delta F = \begin{bmatrix} & & 0 & \\ -\Delta A_0 & \cdots & & -\Delta A_{m-1} \end{bmatrix}, \qquad \Delta G = \begin{bmatrix} 0 & \\ & \Delta A_m \end{bmatrix}.
$$

As $\Delta A_i = D\Theta E_i$, we have

$$
\Delta F - \lambda \Delta G = \widetilde{D}\Theta\,[\,E_0, \quad \ldots, \quad E_{m-1} + \lambda E_m\,] \quad \text{with} \quad \widetilde{D} = -\begin{bmatrix} 0 \\ \vdots \\ 0 \\ D \end{bmatrix}.
$$

Then, using the identity $\det(I + AB) = \det(I + BA)$, valid whenever both AB and BA are defined [47, p. 54],

$$
\begin{aligned}
\det(P(\lambda) + \Delta P(\lambda)) = 0 &\Leftrightarrow \det(F - \lambda G + \Delta F - \lambda \Delta G) = 0 \\
&\Leftrightarrow \det(I + (F - \lambda G)^{-1}(\Delta F - \lambda \Delta G)) = 0 \\
&\Leftrightarrow \det(I + (F - \lambda G)^{-1}\widetilde{D}\Theta\,[\,E_0, \quad \ldots, \quad E_{m-1} + \lambda E_m\,]) = 0 \\
&\Leftrightarrow \det(I + \Theta\,[\,E_0, \quad \ldots, \quad E_{m-1} + \lambda E_m\,]\,(F - \lambda G)^{-1}\widetilde{D}) = 0.
\end{aligned}
$$

Let $M = [\,E_0, \quad \ldots, \quad E_{m-1} + \lambda E_m\,]\,(F - \lambda G)^{-1}\widetilde{D} \in \mathbb{C}^{t \times s}$. Then, using [45, Lem. 1], we have

$$
\eta(\lambda; D, E) = \min_{\Theta \in \mathbb{C}^{s \times t}} \big\{ \|\Theta\| : \det(I + \Theta M) = 0 \big\} = \|M\|^{-1}.
$$

But it is easily verified that

$$
(F - \lambda G) \begin{bmatrix} P(\lambda)^{-1}D \\ \vdots \\ \lambda^{m-1}P(\lambda)^{-1}D \end{bmatrix} = \widetilde{D},
$$

so that $M = E(\lambda)P(\lambda)^{-1}D$.   □

We define the structured ϵ-pseudospectrum by

$$\Lambda_\epsilon(P; D, E) = \{ \lambda \in \mathbb{C} : (P(\lambda) + D\Theta E(\lambda))x = 0 \quad \text{for some } x \neq 0, \|\Theta\| \leq \epsilon \}.$$

Analogously to the unstructured case, $\Lambda_\epsilon(P; D, E) = \{ \lambda \in \mathbb{C} : \eta(\lambda; D, E) \leq \epsilon \}$, and so from Lemma 2.4 we have

$$(2.17) \qquad \Lambda_\epsilon(P; D, E) = \{ \lambda \in \mathbb{C} : \|E(\lambda)P(\lambda)^{-1}D\| \geq \epsilon^{-1} \},$$

which is a generalization of a result of Hinrichsen and Kelb [17, Lem. 2.2] for the ϵ-pseudospectrum of a single matrix.

2.4. Connection between backward error and stability radius. In many mathematical models (e.g., those of a dynamical system) it is required for stability that a matrix has all its eigenvalues in a given open subset $\mathbb{C}_g \neq \emptyset$ of the complex plane. Various stability radii have been defined that measure the ability of a matrix to preserve its stability under perturbations.

We partition the complex plane $\mathbb{C}$ into two disjoint subsets $\mathbb{C}_g$ and $\mathbb{C}_b$, with

$$(2.18) \qquad \mathbb{C} = \mathbb{C}_g \cup \mathbb{C}_b, \quad \mathbb{C}_g \neq \emptyset \text{ an open set.}$$

Consider perturbations of the form in (2.10). Following Pappas and Hinrichsen [30] and Genin and Van Dooren [7], we define the complex structured stability radius of the λ-matrix P with respect to the perturbation structure (D, E) and the partition (2.18) by

$$r_{\mathbb{C}}(P; D, E) = \inf_{\Theta \in \mathbb{C}^{s \times t}} \{ \|\Theta\| : \Lambda(P(\lambda) + \Delta P(\lambda)) \cap \mathbb{C}_b \neq \emptyset, \ \Delta P(\lambda) = D\Theta E(\lambda) \}.$$

Let $\partial\mathbb{C}_b$ be the boundary of $\mathbb{C}_b$. By continuity, we have

$$\begin{aligned}
r_{\mathbb{C}}(P; D, E) &= \inf_{\Theta \in \mathbb{C}^{s \times t}} \{ \|\Theta\| : \Lambda(P(\lambda) + D\Theta E(\lambda)) \cap \partial\mathbb{C}_b \neq \emptyset \} \\
&= \inf_{\lambda \in \partial\mathbb{C}_b} \inf_{\Theta \in \mathbb{C}^{s \times t}} \{ \|\Theta\| : \det(P(\lambda) + D\Theta E(\lambda)) = 0 \} \\
&= \inf_{\lambda \in \partial\mathbb{C}_b} \inf_{x \neq 0} \inf_{\Theta \in \mathbb{C}^{s \times t}} \{ \|\Theta\| : (P(\lambda) + D\Theta E(\lambda))x = 0 \} \\
&= \inf_{\lambda \in \partial\mathbb{C}_b} \eta(\lambda; D, E).
\end{aligned}$$

Thus we have expressed the stability radius as an infimum of the eigenvalue backward error. Using Lemma 2.4 we obtain the following result.

LEMMA 2.5. *If λ is not an eigenvalue of P then*

$$r_{\mathbb{C}}(P; D, E) = \inf_{\lambda \in \partial\mathbb{C}_b} \|E(\lambda)P(\lambda)^{-1}D\|^{-1}$$

and for unstructured perturbations and the p-norm we have

$$r_{\mathbb{C}}(P; I_n, I_{(m+1)n}) = \inf_{\lambda \in \partial\mathbb{C}_b} \left(\|[1 \ \lambda \ \dots \lambda^m]\|_p \|P(\lambda)^{-1}\|_p \right)^{-1}.$$

The result for the unstructured case in the second part of this lemma is also obtained by Pappas and Hinrichsen [30, Cor. 2.4] and Genin and Van Dooren [7, Thm. 2].

3. Computation of pseudospectra. In this section, we consider the computation of $\Lambda_\epsilon(P)$, concentrating mainly on the 2-norm. We develop methods for unstructured perturbations and show how they can be extended to structured perturbations of the form in (2.10).

Lemma 2.1 shows that the boundary of $\Lambda_\epsilon(P)$ comprises points z for which the scaled resolvent norm $p(|z|)\|P(z)^{-1}\|$ equals ϵ^{-1}. Hence, as for pseudospectra of a single matrix, we can obtain a graphical representation of the pseudospectra of a polynomial eigenvalue problem by evaluating the scaled resolvent norm on a grid of points z in the complex plane and sending the results to a contour plotter. We refer to Trefethen [42] for a survey of the state of the art in computation of pseudospectra of a single matrix.

The region of interest in the complex plane will usually be determined by the underlying application or by prior knowledge of the spectrum of P. In the absence of such information we can select a region guaranteed to enclose the spectrum. If A_m is nonsingular (so that all eigenvalues are finite) then by applying the result "$\max_j |\lambda_j(A)| \leq \|A\|$" to the companion form (2.16) we deduce that

$$\max_j |\lambda_j(P)| \leq 1 + \sum_{j=0}^{m-1} \|A_m^{-1} A_j\|_p$$

for any p-norm. Alternatively, we could bound $\max_j |\lambda_j(P)|$ by the largest absolute value of a point in the numerical range of P [24], but computation of this number is itself a nontrivial problem. For much more on bounding the eigenvalues of matrix polynomials see [15].

For the 2-norm, $\|P(z)^{-1}\|_2 = (\sigma_{\min}(P(z)))^{-1}$, where $\sigma_{\min}$ denotes the smallest singular value. If the grid is $\nu \times \nu$ and $\sigma_{\min}$ is computed using the Golub–Reinsch SVD algorithm then the whole computation requires roughly

$$(3.1) \qquad\qquad\qquad \nu^2(8n^3/3 + n^2 m) \text{ flops,}$$

which is prohibitively expensive for matrices of large dimension and a fine grid. Using the fact that $\sigma_{\min}(P(z))$ is the square root of $\lambda_{\min}(P(z)^* P(z))$, we can approximate $\|P(z)^{-1}\|_2$ with the power iteration or Lanczos iteration applied to $P(z)^{-1} P(z)^{-*}$. In the case of a single matrix, Lui [25] introduced the idea of using the Schur form of A in order to speed up the computation of $\lambda_{\min}((A - zI)^*(A - zI))$. Unfortunately, for matrix polynomials of degree $m \geq 2$ no analogue of the Schur form exists (that is, at most two general matrices can be simultaneously reduced to triangular form). We therefore look for other ways to efficiently evaluate or approximate $\|P(z)^{-1}\|$ for many different z.

3.1. Transfer function approach. The idea of writing pseudospectra in terms of transfer functions is not new. Simoncini and Gallopoulos [34] used a transfer function framework to rewrite most of the techniques used to approximate ϵ-pseudospectra of large matrices, yielding interesting comparisons as well as better understanding of the techniques. Hinrichsen and Kelb [17] investigated structured pseudospectra of a single matrix with perturbations of the form in (2.10), and they expressed the structured ϵ-pseudospectrum in terms of a transfer function.

Consider the equation

$$P(z)v = (z^m A_m + z^{m-1} A_{m-1} + \cdots + A_0)v = u.$$

It can be rewritten as

$$(F - zG) \begin{bmatrix} v \\ w_2 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -u \end{bmatrix},$$

where F and G are defined in (2.16). Hence

$$P(z)^{-1}u = v = \begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix} (F - zG)^{-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -u \end{bmatrix} = \begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix} (F - zG)^{-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -I \end{bmatrix} u.$$

Since this equation holds for all u, it follows that

$$(3.2) \qquad P(z)^{-1} = \begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix} (F - zG)^{-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -I \end{bmatrix}.$$

This equality can also be deduced from the theory of λ-matrices [21, Thm. 14.2.1]. We have thus expressed the resolvent in terms of a transfer function.

In control theory, $P(z)^{-1}$ corresponds to the transfer function of the linear time-invariant multivariate system described by

$$G\dot{x}(t) = Fx(t) + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \end{bmatrix} u(t),$$

$$y(t) = \begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix} x(t).$$

Several algorithms have been proposed in the literature [22], [27] to compute transfer functions at a large number of frequencies, most of them assuming that $G = I$. Our objective is to efficiently compute the norm of the transfer function, rather than to compute the transfer function itself.

For structured perturbations we see from (2.17) that the transfer function $P(z)^{-1}$ is replaced by

$$E(z)P(z)^{-1}D = \begin{bmatrix} E(z) & 0 & \cdots & 0 \end{bmatrix} (F - zG)^{-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -D \end{bmatrix}.$$

All the methods described below for the dense case are directly applicable with obvious changes.

We would like a factorization of $F - zG$ that enables efficient evaluation or application of $(F - zG)^{-1}$ for many different z. There are various possibilities, including, when G is nonsingular,

$$F - zG = G(G^{-1}F - zI) = G(W^*TW - zI) = GW^*(T - zI)W,$$

where $G^{-1}F = W^*TW$ is a Schur decomposition, with W unitary and T upper triangular. However this approach is numerically unstable when G is ill conditioned. A numerically stable reduction is obtained by computing the generalized Schur decomposition

$$(3.3) \qquad W^*FZ = T, \quad W^*GZ = S,$$

where W and Z are unitary and T and S are upper triangular. Then

$$(3.4) \qquad P(z)^{-1} = \begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix} Z\,(T - zS)^{-1}\,W^* \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -I \end{bmatrix}.$$

Hence once the generalized Schur decomposition has been computed, we can compute $P(z)^{-1}x$ and $P(z)^{-*}x$ at a cost of $O((mn)^2)$ flops, since $T - zS$ is triangular of dimension mn. For the 2-norm we can therefore efficiently approximate $\|P(z)^{-1}\|$ using inverse iteration or the inverse Lanczos iteration, that is, the power method or the Lanczos method applied to $P(z)^{-1}P(z)^{-*}$.

The cost of the computation breaks into two parts: the cost of the initial transformations and the cost of the computations at each of the ν^2 grid points. Assuming that (3.3) is computed using the QZ algorithm [9, Sec. 7.7] and the average number of power method or Lanczos iterations per grid point is k, the total cost is about

$$66(mn)^3 + k\nu^2(8mn^2 + 3(mn)^2) \text{ flops.}$$

For the important special case $m = 2$ (the quadratic eigenvalue problem), this cost is

$$(3.5) \qquad 528n^3 + 28k\nu^2n^2 \text{ flops.}$$

Comparing with (3.1) we see that this method is a significant improvement over the SVD-based approach for a sufficiently fine grid and a small degree m.

For the 2-norm note that, because of the two outer factors in (3.4), we cannot discard the unitary matrices Z and W, unlike in the analogous expression for the resolvent of a single matrix in the standard eigenproblem. For the 1- and ∞-norms we can efficiently estimate $\|P(z)^{-1}\|$ using the algorithm of Higham and Tisseur [14], which requires only the ability to multiply matrices by $P(z)^{-1}$ and $P(z)^{-*}$.

An alternative to the generalized Schur decomposition is the generalized Hessenberg-triangular form, which differs from (3.3) in that one of T and S is upper Hessenberg. The Hessenberg form is cheaper to compute but more expensive to work with. It leads to a smaller overall flop count when $k\nu^2 \gtrsim 25mn$.

3.2. Factorizing the quadratic polynomial. The transfer function-based method of the previous section has the drawback that it factorizes matrices of dimension m times those of the original polynomial matrix. We now describe another method, particular to the quadratic case, that does not increase the size of the problem.

Suppose we can find a matrix S such that $A_2S^2 + A_1S + A_0 = 0$, that is, a solvent of the quadratic matrix equation $A_2X^2 + A_1X + A_0 = 0$. Then

$$(3.6) \qquad Q(z) := z^2A_2 + zA_1 + A_0 = -(A_1 + A_2S + zA_2)(S - zI).$$

If we compute the Schur decomposition

$$S = QTQ^*$$

and the generalized Schur decomposition

$$W^*(A_1 + A_2 S)Z = R_1, \quad W^* A_2 Z = R_2$$

then

$$(3.7) \qquad Q(z)^{-1} = -Q(T - zI)^{-1} Q^* Z (R_1 + zR_2)^{-1} W^*,$$

so a vector can be premultiplied by $Q(z)^{-1}$ or its conjugate transpose in $O(n^2)$ flops for any z. Moreover, for the 2-norm we can drop the outer Q and W^* factors in (3.7), by unitary invariance, and hence we do not need to form W. For the 2-norm, the total cost of this method is

$$(3.8) \qquad c_S + 77n^3 + 10k\nu^2 n^2 \text{ flops},$$

where c_S is the cost of computing a solvent and we have assumed that we precompute $Q^* Z$. Comparing this flop count with (3.5) we see that the cost per grid point of the solvent approach is much lower.

The success of this method depends on two things: the existence of solvents and being able to compute one at a reasonable cost. Some sufficient conditions for the existence of a solvent are summarized in [13]. In particular, for an overdamped problem, one for which A_2 and A_1 are Hermitian positive definite, A_0 is Hermitian positive semidefinite, and $(x^* A_1 x)^2 > 4(x^* A_2 x)(x^* A_0 x)$ for all $x \neq 0$, a solvent is guaranteed to exist.

Various methods are available for computing solvents [12], [13]. One of the most generally useful is Newton's method, optionally with exact line searches, which requires a generalized Sylvester equation in $n \times n$ matrices to be solved on each iteration, at a total cost of about $56n^3$ flops per iteration. If Newton's method converges within 8 iterations or so, so that $c_S \leq 448n^3$ flops, this approach is certainly competitive in cost with the transfer function approach.

When there is a gap between the n largest and n smallest eigenvalues ordered by modulus, as is the case for overdamped problems [20, Sec. 7.6], Bernoulli iteration is an efficient way of computing the dominant or minimal solvent S [13]. If t iterations are needed for convergence to the dominant or minimal solvent then the cost of Bernoulli iteration is about $c_S = 4tn^3$ flops. Bernoulli iteration converges only linearly, but convergence is fast if the eigenvalue gap is large.

A third approach to computing a solvent is to use a Schur method from [13], based on the following theorem. Let F and G be defined as in (2.16), so that

$$F = \begin{bmatrix} 0 & I \\ -A_0 & -A_1 \end{bmatrix}, \qquad G = \begin{bmatrix} I & 0 \\ 0 & A_2 \end{bmatrix}.$$

THEOREM 3.1 (Higham and Kim [13]). *All solvents of $Q(X)$ are of the form* $X = Z_{21}Z_{11}^{-1} = Q_{11}T_{11}S_{11}^{-1}Q_{11}^{-1}$, *where*

$$(3.9) \qquad Q^* FZ = T, \quad Q^* GZ = S$$

is a generalized Schur decomposition with Q and Z unitary and T and S upper triangular, and where all matrices are partitioned as block 2×2 matrices with $n \times n$ blocks.

The method consists of computing the generalized Schur decomposition (3.9) by the QZ algorithm and then forming $S = Z_{21}Z_{11}^{-1}$. The generalized Schur decomposition may need to be reordered in order to obtain a nonsingular Z_{11}. Note that the

unitary factor Q does not need to be formed. For this method, $c_S = 50(2n)^3 + rn^2$, where the constant r depends on the amount of reordering required. From (3.8), the total cost is now

$$479n^3 + rn^2 + 10k\nu^2 n^2 \text{ flops,}$$

which is much more favorable than the cost (3.5) of the transfer function method.

For higher degree polynomials we can generalize this approach by attempting to factorize P into linear factors by recursively computing solvents. However, for degrees m greater than 2 classes of problem for which a factorization into linear factors exists are less easily identified and the cost of Newton's method (for example) is much higher than for $m = 2$ [19].

3.3. Large-scale computation. All the methods described above are intended for small- to medium-scale problems for which Schur and other reductions are possible. For large, possibly sparse, problems, different techniques are necessary. These techniques can be classified into two categories: those that project to reduce the size of the problem and then compute the pseudospectra of the reduced problem, and those that approximate the norm of the resolvent directly.

3.3.1. Projection approach. For a single matrix, A, Toh and Trefethen [39] and Wright and Trefethen [48] approximate the resolvent norm by the Arnoldi method; that is, they approximate $\|(A - zI)^{-1}\|_2$ by $\|(H_m - zI)^{-1}\|_2$ or by $\sigma_{\min}(\widetilde{H}_m - z\widetilde{I})$, where H_m is the square Hessenberg matrix of dimension $m \ll n$ obtained from the Arnoldi process and $\widetilde{H}_m$ is the matrix H_m augmented by an extra row. Simoncini and Gallopoulos [34] show that a better but more costly approximation is obtained by approximating $\|(A - zI)^{-1}\|_2$ with $\|V_m^*(A - zI)^{-1}V_{m+1}\|_2$, where V_m is the orthonormal basis generated during the Arnoldi process. These techniques are not applicable to the polynomial eigenvalue problem of degree larger than one because of the lack of a Schur form for the Arnoldi method to approximate.

A way of approximating $\|P(z)^{-1}\|$ for all z is through a projection of $P(z)^{-1}$ onto a lower dimensional subspace. Let V_k be an $n \times k$ matrix with orthonormal columns. We can apply one of the techniques described in the previous sections to compute pseudospectra of the projected polynomial eigenvalue problem $\widetilde{P}(\lambda) = V_k^* P(\lambda) V_k$. A possible choice for V_k is an orthonormal basis of k selected linearly independent eigenvectors of $P(\lambda)$. In this case, $\widetilde{P}(\lambda)$ is the matrix representation of the projection of $P(\lambda)$ onto the subspace spanned by the selected eigenvectors. The eigenvectors can be chosen to correspond to parts of the spectrum of interest and can be computed using the Arnoldi process on the companion form pencil (F, G) or directly on $P(\lambda)$ with the Jacobi–Davidson method or its variants [26], [35]. In the latter case, the matrix V_k is built during the Davidson process.

3.3.2. Direct approach. This approach consists of approximating $\|P(z)^{-1}\|$ at each grid point z. Techniques analogous to those used for single matrices can be applied, such as the Lanczos method applied to $P(z)^*P(z)$ or its inverse. We refer the reader to [42] for more details and further references.

4. Applications and numerical experiments. We give a selection of applications of pseudospectra for polynomial eigenvalue problems, using them to illustrate the performance of our methods for computing pseudospectra. All our examples are for 2-norm pseudospectra.

4.1. The wing problem. The first example is based on a quadratic polynomial $Q(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0$ from [6, Sec. 10.11], with numerical values modified as in [20, Sec. 5.3]. The eigenproblem for $Q(\lambda)$ arose from the analysis of the oscillations of a wing in an airstream. The matrices are

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

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

The left plot in Figure 4.1 shows the boundaries of ϵ-pseudospectra with perturbations measured in the absolute sense ($\alpha_i \equiv 1$), with ϵ between 10^{-3} and $10^{-0.8}$. The eigenvalues are plotted as dots. Another way of approximating a pseudospectrum is by random perturbations of the original matrices [41]. We generated 200 triples of complex random normal perturbation matrices $(\Delta A_1, \Delta A_2, \Delta A_3)$ with $\|\Delta A_j\|_2 = 10^{-0.8}$, $j = 1{:}3$. In the right plot of Figure 4.1 are superimposed as small dots the eigenvalues of the perturbed polynomials $\lambda^2 (A_2 + \Delta A_2) + \lambda(A_1 + \Delta A_1) + \Delta A_0 + \Delta A_0$. The solid curve marks the boundary of the ϵ-pseudospectrum for $\epsilon = 10^{-0.8}$. Both pictures show that the pair of complex eigenvalues $\lambda = -0.88 \pm 8.4i$ are much more sensitive to perturbations than the other two complex pairs.

The eigenvalues of $Q(\lambda)$ are the same as those of the linearized problem $\mathcal{A} - \lambda I$, where

(4.1)
$$\mathcal{A} = \begin{bmatrix} 0 & I \\ -A_2^{-1}A_0 & -A_2^{-1}A_1 \end{bmatrix}.$$

Figure 4.2 shows boundaries of ϵ-pseudospectra for this matrix, for the same ϵ as in Figure 4.1. Clearly, the ϵ-pseudospectra of the linearized problem (4.1) do not give useful information about the behavior of the eigensystem of $Q(\lambda)$ under perturbations. This emphasizes the importance of defining and computing pseudospectra for the quadratic eigenvalue problem in its original form.

4.2. Mass-spring system. We now consider the connected damped mass-spring system illustrated in Figure 4.3. The ith mass of weight m_i is connected to the $(i+1)$st mass by a spring and a damper with constants k_i and d_i, respectively. The ith mass is also connected to the ground by a spring and a damper with constants κ_i and τ_i, respectively. The vibration of this system is governed by a second-order differential equation

$$M\frac{d^2}{dt^2}x + C\frac{d}{dt}x + Kx = 0,$$

where the mass matrix $M = \mathrm{diag}(m_1, \ldots, m_n)$ is diagonal, and the damping matrix C and stiffness matrix K are symmetric tridiagonal. The differential equation leads to the quadratic eigenvalue problem

(4.2)
$$(\lambda^2 M + \lambda C + K)x = 0.$$

In our experiments, we took all the springs (respectively, dampers) to have the same constant $\kappa = 5$ (respectively, $\tau = 10$), except the first and last, for which the constant is 2κ (respectively, 2τ), and we took $m_i \equiv 1$. Then

$$C = \tau \, \mathrm{tridiag}(-1, 3, -1), \quad K = \kappa \, \mathrm{tridiag}(-1, 3, -1),$$
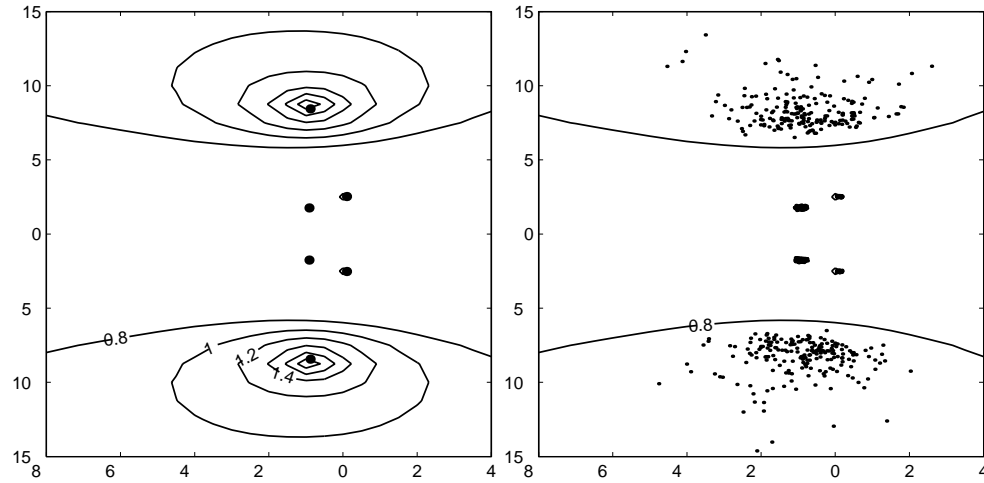
FIG. 4.1. *Wing problem. Left: $\Lambda_\epsilon(Q)$, for $\epsilon \in [10^{-3}, 10^{-0.8}]$. Right: approximation to ϵ-pseudospectrum with $\epsilon = 10^{-0.8}$.*
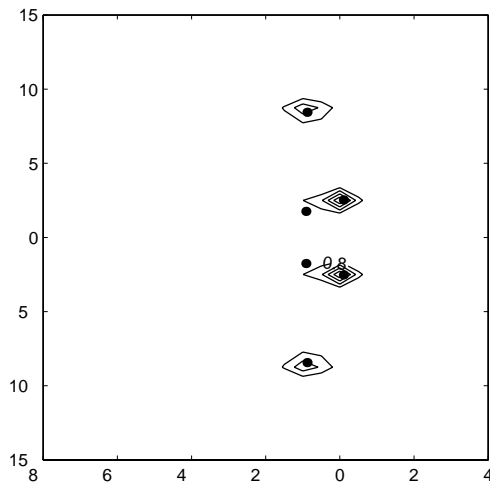

FIG. 4.2. *Wing problem. $\Lambda_\epsilon(\mathcal{A})$, for $\mathcal{A}$ in (4.1) with $\epsilon \in [10^{-3}, 10^{-0.8}]$.*

and the quadratic eigenvalue problem is overdamped. We take an $n = 250$ degree of freedom mass-spring system over a 100×100 grid. A plot of the pseudospectra is given in Figure 4.4.

For this problem we compare all the methods described. In the solvent approach exact line searches were used in Newton's method and no reordering was used in the generalized Schur method. The solvents from the Bernoulli and Schur methods were refined by one step of Newton's method. The Bernoulli iteration converged in 12 iterations while only 6 iterations were necessary for Newton's method. The Lanczos inverse iteration converged after 3 iterations on average. In Table 4.1 we give the estimated flop counts, using the formulae from section 3, together with execution times. The computations were performed in MATLAB 6, which is an excellent environment for investigating pseudospectra. While the precise times are not important, the con-
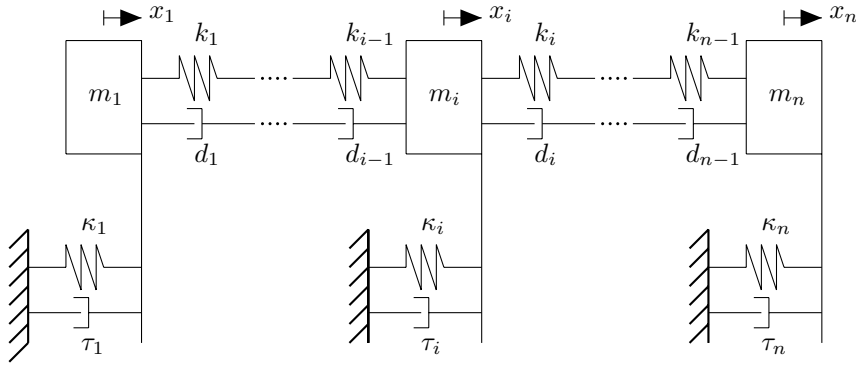
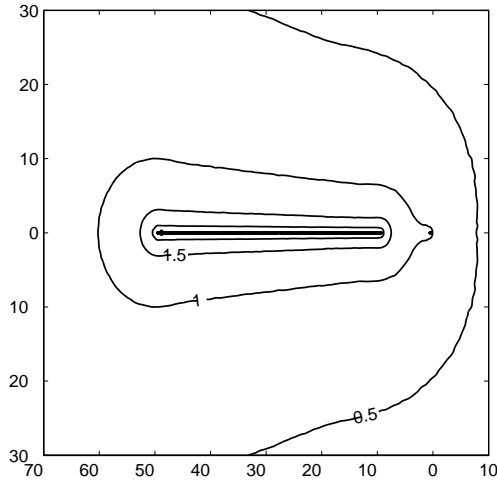FIG. 4.3. *An n degree of freedom damped mass-spring system.*

FIG. 4.4. *Pseudospectra of a* 250 *degree of freedom damped mass-spring system on a* 100×100 *grid.*

clusion is clear: in this example, the three solvent-based methods are much faster than the SVD and transfer function methods. (The high speed of the SVD method relative to its flop count is attributable to MATLAB's very efficient `svd` function.)

4.3. Acoustic problem. Acoustic problems with damping can give rise to large quadratic eigenvalue problems (4.2), where, again, M is the mass matrix, C is the damping matrix, and K the stiffness matrix. We give in Figure 4.5 the sparsity pattern of the three matrices M, C, and K of order 107 arising from a model of a speaker box [1]. These matrices are symmetric and the sparsity patterns of M and K are identical. There is a large variation in the norms: $\|M\|_2 = 1$, $\|C\|_2 = 0.06$, $\|K\|_2 = 9.9 \times 10^6$.

We plot in Figure 4.6 pseudospectra with perturbations measured in both an absolute sense ($\alpha_1 = \alpha_2 = \alpha_3 = 1$) and a relative sense ($\alpha_1 = \|M\|_2$, $\alpha_2 = \|C\|_2$, $\alpha_3 = \|K\|_2$), together with pseudospectra of the corresponding standard eigenvalue problem of the form (4.1). The eigenvalues are all pure imaginary and are marked by dots on the plot. The two first plots are similar, both showing that the most sensitive

TABLE 4.1
Comparison in terms of flops and execution time of different techniques.

Method	Estimated cost in flops	Execution time
Golub–Reinsch SVD	$26747n^3$	102 min
Transfer function	$3408n^3$	106 min
Solvent: Newton	$1589n^3$	39 min
Solvent: Bernoulli	$1325n^3$	36 min
Solvent: Schur	$1677n^3$	37 min
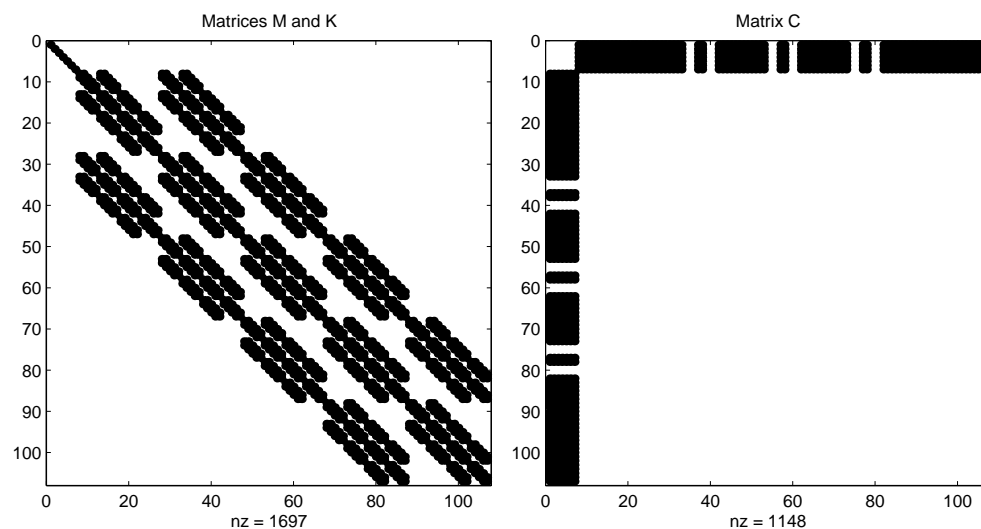
FIG. 4.5. *Sparsity patterns of the three 107×107 matrices $M, C,$ and K of an acoustic problem.*

eigenvalues are located at the extremities of the spectrum; the contour lines differ mainly around the zero eigenvalue. The last plot is very different; clearly it is the eigenvalues close to zero that are the most sensitive to perturbations of the standard eigenproblem form.

We mention that for this problem we have been unable to compute a solvent.

4.4. Closed loop system. In multi-input and multioutput systems in control theory the location of the eigenvalues of matrix polynomials determine the stability of the system. Figure 4.7 shows a closed-loop system with feedback with gains 1 and $1 + \alpha$, $\alpha > 0$. The associated matrix polynomial is given by

$$P(z) = z^2 I + z \begin{bmatrix} 0 & 1 + \alpha \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 1/2 & 0 \\ 0 & 1/4 \end{bmatrix}.$$

We are interested in the values of α for which $P(z)$ has all its eigenvalues inside the unit circle. By direct calculation with $\det(P(z))$, using the Routh array, for example, it can be shown that $P(z)$ has all its eigenvalues inside the unit circle if and only if $\alpha < 0.875$.

The matrix $P(z)$ can be viewed as a perturbed matrix polynomial with structured perturbations:
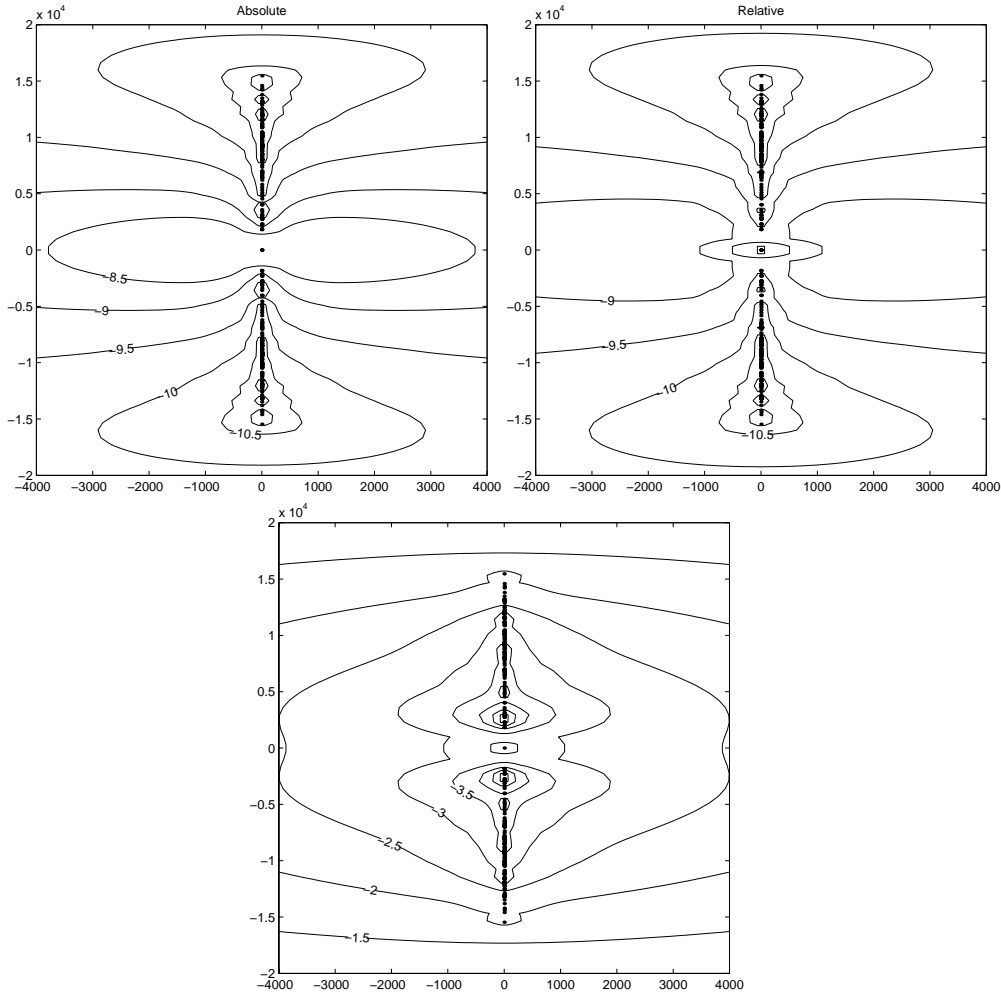
$$P(z) = P(z) \mid_{\alpha=0} + D\Theta E(z),$$

FIG. 4.6. *Acoustic problem, $n = 107$, 70×70 grid. Perturbations measured in an absolute sense* (*top left*) *and relative sense* (*top right*). *Pseudospectra of the equivalent standard eigenvalue problem are shown at the bottom.*

where

$$D = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \Theta = \alpha, \qquad E(z) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} I \\ zI \\ z^2 I \end{bmatrix}.$$

We show in Figure 4.8 the structured pseudospectra as defined by (2.17). The dashed lines mark the unit circle. Since the outermost contour has value $\alpha = 0.875$ and just touches the unit circle, this picture confirms the value for the maximal α that we obtained analytically.

4.5. The Orr–Sommerfeld equation. The Orr–Sommerfeld equation is a linearization of the incompressible Navier–Stokes equations in which the perturbations in velocity and pressure are assumed to take the form

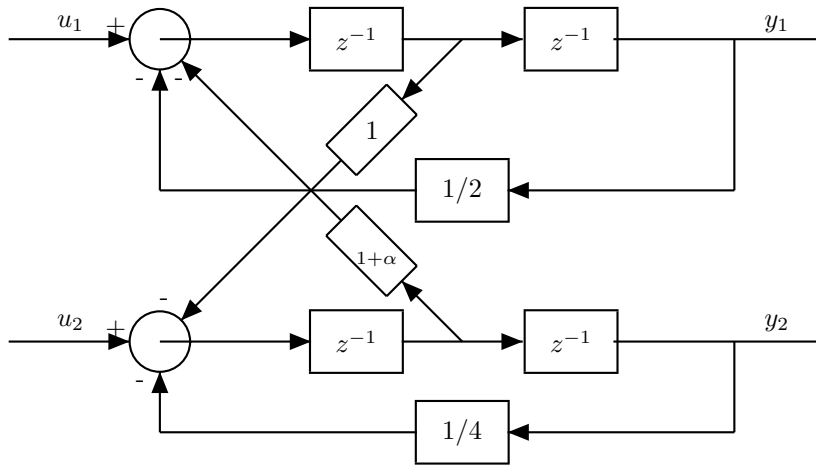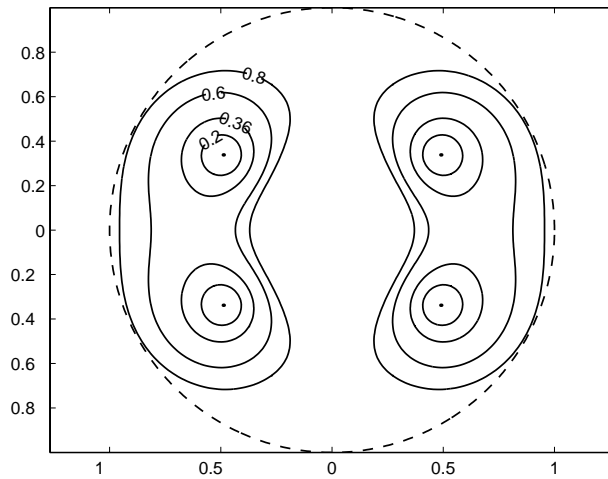$$\Phi(x, y, t) = \phi(y)e^{i(\lambda x - \omega t)},$$

FIG. 4.7. *Closed-loop system with feedback gains* 1 *and* $1 + \alpha$.

FIG. 4.8. *Structured pseudospectra of a closed-loop system with one-parameter feedback.*

where λ is a wavenumber and ω is a radian frequency. For a given Reynolds number R, the Orr–Sommerfeld equation may be written

$$(4.3) \qquad \left[\left(\frac{d^2}{dy^2} - \lambda^2 \right)^2 - iR \left\{ (\lambda U - \omega) \left(\frac{d^2}{dy^2} - \lambda^2 \right) - \lambda U'' \right\} \right] \phi = 0.$$

We consider plane Poiseuille flow between walls at $y = \pm 1$ and with velocity $U(y) = 1 - y^2$ in the streamwise x direction, for which the boundary conditions are

$$\phi(\pm 1) = 0, \quad \phi'(\pm 1) = 0.$$

For a given real value of R, the boundary conditions will be satisfied only for certain combinations of values of λ and ω. Two cases are of interest.

Case 1. *Temporal stability.* If λ is fixed and real, then (4.3) is linear in the parameter ω and corresponds to a generalized eigenvalue problem. The perturbations

are periodic in x and grow or decay in time depending on the sign of the imaginary part of ω. This case has been studied with the help of pseudospectra by Reddy, Schmid, and Henningson [32].

Case 2. *Spatial stability.* For most real flows, the perturbations are periodic in time, which means that ω is real. Then the sign of the imaginary part of λ determines whether the perturbations will grow or decay in space. In this case, the parameter is λ, which appears to the fourth power in (4.3), so we obtain a quartic polynomial eigenvalue problem. Bridges and Morris [2] calculated the spectrum of (4.3) using a finite Chebyshev series expansion of ϕ combined with the Lanczos tau method and they computed the spectrum of the quartic polynomial by two methods: the QR algorithm applied to the corresponding standard eigenvalue problem in companion form, and Bernoulli iteration applied to determine a minimal solvent and hence to obtain the n eigenvalues of minimal modulus.

For our estimation of the pseudospectra of the Orr–Sommerfeld equation we use a Chebyshev spectral discretization that combines an expansion in Chebyshev polynomials and collocation at the Chebyshev points with explicit enforcement of the boundary conditions. We are interested in the eigenvalues λ that are the closest to the real axis, and we need $\mathrm{Im}(\lambda) > 0$ for stability. The linear eigenvalue problem (Case 1) has been solved by Orszag [29]. The critical neutral point corresponding to λ and ω both real for minimum R was found at $R = 5772$ and $\lambda = 1.02056$ with the frequency $\omega = 0.26943$ [2], [29]. For our calculations we set R and ω to these values and we computed the modes λ, taking $N = 64$, which gives matrices of order $N - 1$. The first few modes are plotted in Figure 4.9. For the first mode we obtained $\lambda = 1.02056 + 9.7 \times 10^{-7}i$, which compares favorably with the result of Orszag. Figure 4.10 shows the pseudospectra in a region around the first few modes on a 100×100 grid, with $\alpha_i = \|A_i\|_2$ except that $\alpha_4 = 0$, since A_4 is the identity matrix and is not subject to uncertainty. The plot shows that the first mode is very sensitive. Interestingly, the second and subsequent modes are almost as sensitive, with perturbations of order 10^{-9} in the matrix coefficients being sufficient to move all these modes across the real axis, making the flow unstable. The pseudospectra thus give a guide to the accuracy with which computations must be carried out for the numerical approximations to the modes to correctly determine the location of the modes. For more on the interpretation of pseudospectra for this problem, see [32] and [44].

Again, for comparison we computed the pseudospectra of the corresponding standard eigenvalue problem. The picture was qualitatively similar, but the contour levels were several orders of magnitude smaller, thus not revealing the true sensitivity of the problem.
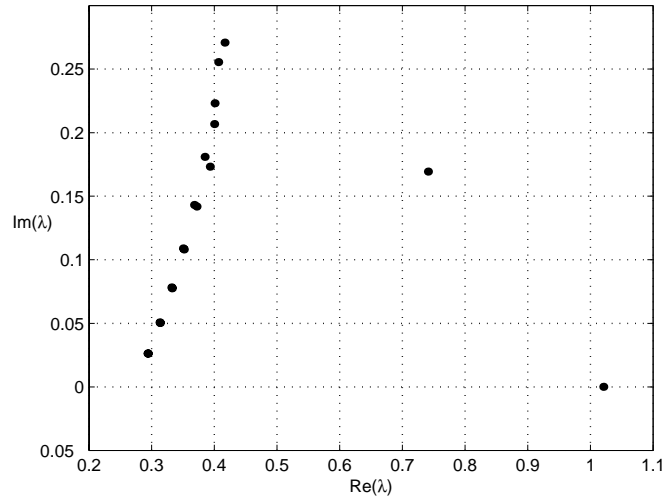
Fig. 4.9. *The first few modes of the spectrum of the Orr–Sommerfeld equation for $R = 5572$ and $\omega = 0.26943$.*
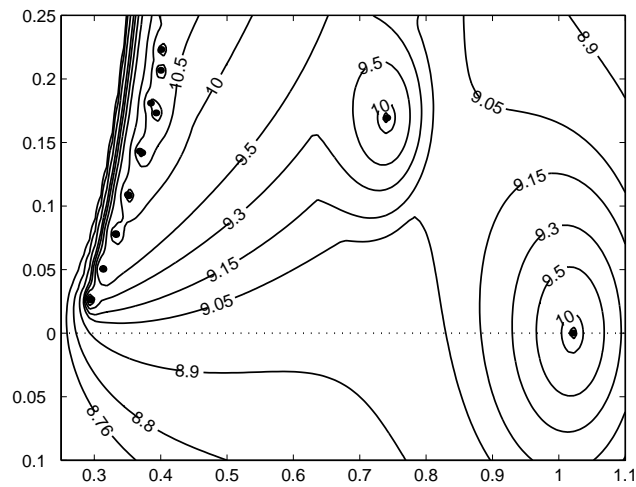

Fig. 4.10. *Pseudospectra of the Orr–Sommerfeld equation for $R = 5572$ and $\omega = 0.26943$.*

REFERENCES

[1] Z. Bai, *private communication*, 1999.

[2] T. J. Bridges and P. J. Morris, *Differential eigenvalue problems in which the parameter appears nonlinearly*, J. Comput. Phys., 55 (1984), pp. 437–460.

[3] M. Embree and L. N. Trefethen, *Pseudospectra gateway*, http://www.comlab.ox.ac.uk/pseudospectra/.

[4] V. Frayssé, M. Gueury, F. Nicoud, and V. Toumazou, *Spectral Portraits for Matrix Pencils*, Technical Report TR/PA/96/19, CERFACS, Toulouse, France, 1996.

[5] V. Frayssé and V. Toumazou, *A note on the normwise perturbation theory for the regular generalized eigenproblem*, Numer. Linear Algebra Appl., 5 (1998), pp. 1–10.

[6] R. A. Frazer, W. J. Duncan, and A. R. Collar, *Elementary Matrices and Some Applications to Dynamics and Differential Equations*, 10th ed., Cambridge University Press, 1938, 1963 printing.

[7] Y. Genin and P. M. Van Dooren, *Stability Radii of Polynomial Matrices*, manuscript, 1999.

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

[9] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.

[10] D. J. Higham and N. J. Higham, *Structured backward error and condition of generalized eigenvalue problems*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 493–512.

[11] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, 1996.

[12] N. J. Higham and H.-M. Kim, *Solving a Quadratic Matrix Equation by Newton's Method with Exact Line Searches*, Numerical Analysis Report 339, Manchester Centre for Computational Mathematics, Manchester, 1999; SIAM J. Matrix Anal. Appl., to appear.

[13] N. J. Higham and H.-M. Kim, *Numerical analysis of a quadratic matrix equation*, IMA J. Numer. Anal., 20 (2000), pp. 499–519.

[14] N. J. Higham and F. Tisseur, *A block algorithm for matrix 1-norm estimation, with an application to 1-norm pseudospectra*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1185–1201.

[15] N. J. Higham and F. Tisseur, *Bounds for Eigenvalues of Matrix Polynomials*, Numerical Analysis Report 371, Manchester Centre for Computational Mathematics, Manchester, 2001; Linear Algebra Appl., to appear.

[16] N. J. Higham and F. Tisseur, *More on Pseudospectra for Polynomial Eigenvalue Problems and Applications in Control Theory*, Numerical Analysis Report 372, Manchester Centre for Computational Mathematics, Manchester, 2001.

[17] D. Hinrichsen and B. Kelb, *Spectral value sets: A graphical tool for robustness analysis*, Systems Control Lett., 21 (1993), pp. 127–136.

[18] D. Hinrichsen and A. J. Pritchard, *Real and complex stability radii: A survey*, in Control of Uncertain Systems, D. Hinrichsen and B. Mårtensson, eds., Progr. Systems Control Theory 6, Birkhäuser, Boston, 1990, pp. 119–162.

[19] W. Kratz and E. Stickel, *Numerical solution of matrix polynomial equations by Newton's method*, IMA J. Numer. Anal., 7 (1987), pp. 355–369.

[20] P. Lancaster, *Lambda-Matrices and Vibrating Systems*, Pergamon Press, Oxford, 1966.

[21] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed., Academic Press, London, 1985.

[22] A. J. Laub, *Efficient multivariable frequency response computations*, IEEE Trans. Automat. Control, AC-26 (1981), pp. 407–408.

[23] P.-F. Lavallée, *Nouvelles Approches de Calcul du ε-Spectre de Matrices et de Faisceaux de Matrices*, Ph.D. thesis, L'Université de Rennes 1, Rennes, France, 1997.

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

[25] S. H. Lui, *Computation of pseudospectra by continuation*, SIAM J. Sci. Comput., 18 (1997), pp. 565–573.

[26] K. Meerbergen, *Locking and restarting quadratic eigenvalue solvers*, SIAM J. Sci. Comput., 22 (2001), pp. 1814–1839.

[27] P. Misra and R. V. Patel, *A determinant identity and its application in evaluating frequency response matrices*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 248–255.

[28] R. G. Mosier, *Root neighborhoods of a polynomial*, Math. Comp., 47 (1986), pp. 265–273.

[29] S. A. Orszag, *Accurate solution of the Orr–Sommerfeld stability equation*, J. Fluid Mech., 50 (1971), pp. 689–703.

[30] G. Pappas and D. Hinrichsen, *Robust stability of linear systems described by higher order*

dynamic equations, IEEE Trans. Automat. Control, 38 (1993), pp. 1430–1435.

[31] C. R. RAO AND S. K. MITRA, *Generalized Inverse of Matrices and Its Applications*, Wiley, New York, 1971.

[32] S. C. REDDY, P. J. SCHMID, AND D. S. HENNINGSON, *Pseudospectra of the Orr–Sommerfeld operator*, SIAM J. Appl. Math., 53 (1993), pp. 15–47.

[33] K. S. RIEDEL, *Generalized epsilon-pseudospectra*, SIAM J. Numer. Anal., 31 (1994), pp. 1219–1225.

[34] V. SIMONCINI AND E. GALLOPOULOS, *Transfer functions and resolvent norm approximation of large matrices*, Electron. Trans. Numer. Anal., 7 (1998), pp. 190–201.

[35] G. L. G. SLEIJPEN, A. G. L. BOOTEN, D. R. FOKKEMA, AND H. A. VAN DER VORST, *Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems*, BIT, 36 (1996), pp. 595–633.

[36] F. TISSEUR, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl., 309 (2000), pp. 339–361.

[37] F. TISSEUR AND K. MEERBERGEN, *The quadratic eigenvalue problem*, Numerical Analysis Report 370, Manchester Centre for Computational Mathematics, Manchester, 2000; SIAM Rev., 43 (2001), pp. 235–286.

[38] K.-C. TOH AND L. N. TREFETHEN, *Pseudozeros of polynomials and pseudospectra of companion matrices*, Numer. Math., 68 (1994), pp. 403–425.

[39] K.-C. TOH AND L. N. TREFETHEN, *Calculation of pseudospectra by the Arnoldi iteration*, SIAM J. Sci. Comput., 17 (1996), pp. 1–15.

[40] V. TOUMAZOU, *Portraits Spectraux de Matrices: Un Outil d'Analyse de la Stabilité*, Ph.D. thesis, Université Henri Poincaré, Nancy-I, France, 1996.

[41] L. N. TREFETHEN, *Pseudospectra of matrices*, in Numerical Analysis 1991, Proceedings of the 14th Dundee Conference, D. F. Griffiths and G. A. Watson, eds., Pitman Res. Notes Math. Ser. 260, Longman Scientific and Technical, Harlow, UK, 1992, pp. 234–266.

[42] L. N. TREFETHEN, *Computation of pseudospectra*, Acta Numer., 8 (1999), pp. 247–295.

[43] L. N. TREFETHEN, *Spectra and pseudospectra*, in The Graduate Student's Guide to Numerical Analysis '98, M. Ainsworth, J. Levesley, and M. Marletta, eds., Springer-Verlag, Berlin, 1999, pp. 217–250.

[44] L. N. TREFETHEN, A. E. TREFETHEN, S. C. REDDY, AND T. A. DRISCOLL, *Hydrodynamic stability without eigenvalues*, Science, 261 (1993), pp. 578–584.

[45] P. M. VAN DOOREN AND V. VERMAUT, *On stability radii of generalized eigenvalue problems*, in European Control Conference, paper FR-M-H6, 1997.

[46] J. L. M. VAN DORSSELAER, *Pseudospectra for matrix pencils and stability of equilibria*, BIT, 37 (1997), pp. 833–845.

[47] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, 1965.

[48] T. G. WRIGHT AND L. N. TREFETHEN, *Large-scale computation of pseudospectra using ARPACK and* `eigs`, Report 00/11, Numerical Analysis Group, Oxford University Computing Laboratory, Oxford, 2000; SIAM J. Sci. Comput., to appear.