IMA Journal of Numerical Analysis (2005) **25**, 648–669
doi:10.1093/imanum/dri012
Advance Access publication on March 7, 2005

Algorithms for the computation of the pseudospectral radius and the numerical radius of a matrix

EMRE MENGI† AND MICHAEL L. OVERTON‡
Computer Science Department, Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA

[Received on 25 March 2004; revised on 9 November 2004]

Two useful measures of the robust stability of the discrete-time dynamical system $x_{k+1} = Ax_k$ are the ϵ-pseudospectral radius and the numerical radius of A. The ϵ-pseudospectral radius of A is the largest of the moduli of the points in the ϵ-pseudospectrum of A, while the numerical radius is the largest of the moduli of the points in the field of values. We present globally convergent algorithms for computing the ϵ-pseudospectral radius and the numerical radius. For the former algorithm, we discuss conditions under which it is quadratically convergent and provide a detailed accuracy analysis giving conditions under which the algorithm is backward stable. The algorithms are inspired by methods of Byers, Boyd–Balakrishnan, He–Watson and Burke–Lewis–Overton for related problems and depend on computing eigenvalues of symplectic pencils and Hamiltonian matrices.

Keywords: pseudospectrum; field of values; robust stability; ϵ-pseudospectral radius; numerical radius; quadratically convergent; backward stability; singular pencil; Hamiltonian matrix; symplectic pencil.

1. Introduction

The convergence of a discrete-time dynamical system

$$x_k = Ax_{k-1}, \tag{1.1}$$

where $x_k \in \mathbb{C}^n$, $A \in \mathbb{C}^{n\times n}$, depends on the norms of the powers of A. In the asymptotic sense, when $k \to \infty$, eigenvalues provide all the information needed to analyse (1.1). Specifically, it is well known that $\lim_{k\to\infty} \|x_k\| = 0$ is satisfied for all x_0 if and only if A is stable, i.e. all of the eigenvalues of A lie inside the unit circle. Here and throughout this paper $\|\cdot\|$ refers to the 2-norm. Moreover, for generic x_0, the asymptotic decay rate of (1.1) can be measured by the spectral radius of A,

$$\rho(A) = \max\{|\lambda|\colon \lambda \in \Lambda(A)\},$$

where $\Lambda(A)$ is the spectrum (set of eigenvalues) of A. The smaller the spectral radius is, the faster the system (1.1) converges asymptotically.

On the other hand, eigenvalues themselves do not reveal the behavior of (1.1) for finite time unless A is normal. For a non-normal matrix A, even though A is stable, the norm of the kth power $\|A^k\|$ can be arbitrarily large. Two related notions, the field of values and the pseudospectrum, provide additional insight. The field of values of A is

$$F(A) = \{z^*Az\colon z^*z = 1,\ z \in \mathbb{C}^n\}, \tag{1.2}$$

†Email: mengi@cs.nyu.edu
‡Email: overton@cs.nyu.edu

and for real $\epsilon > 0$, the ϵ-pseudospectrum is

$$\Lambda_\epsilon(A) = \{z\colon z \in \Lambda(A+E), \text{ for some } E \text{ such that } \|E\| \leqslant \epsilon\}. \tag{1.3}$$

The site (Embree & Trefethen, 2005) contains detailed information about pseudospectra, including a comprehensive bibliography and software links.

Both the field of values and the ϵ-pseudospectrum of A are compact sets containing the eigenvalues of A. Thus, it makes sense to talk about the points in these sets that are located furthest away from the origin. Both the numerical radius of A

$$r(A) = \max\{|w|\colon w \in F(A)\} \tag{1.4}$$

and the analogous measure defined on the ϵ-pseudospectrum, the ϵ-pseudospectral radius

$$\rho_\epsilon(A) = \max\{|z|\colon z \in \Lambda_\epsilon(A)\}, \tag{1.5}$$

are convenient for estimating the norms of the powers of A. Figure 1 illustrates the pseudospectra of a random matrix (i.e. the real and imaginary components of the entries of the matrix are chosen independently from a normal distribution with mean 0 and standard deviation 1) for various values of ϵ together with the point in the ϵ-pseudospectrum with the largest modulus for $\epsilon = 1$. Similarly, Fig. 2 shows the point where the numerical radius of the same matrix is attained on the field of values.

It follows (Horn & Johnson, 1991) from the definition (1.4) that the numerical radius of A is within factor of 0.5 of $\|A\|$,

$$\frac{\|A\|}{2} \leqslant r(A) \leqslant \|A\|. \tag{1.6}$$

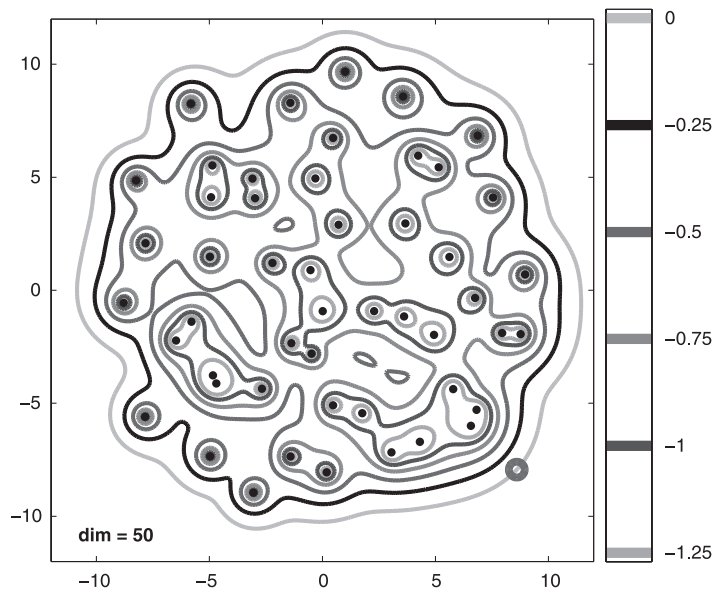

FIG. 1. The eigenvalues (solid dots) and the ϵ-pseudospectra of a random 50×50 matrix are shown for various values of ϵ. The bar on the right shows the values of ϵ on a log 10 scale. The point where the ϵ-pseudospectral radius is attained for $\epsilon = 1$ is in the lower right corner and is marked with a circle.

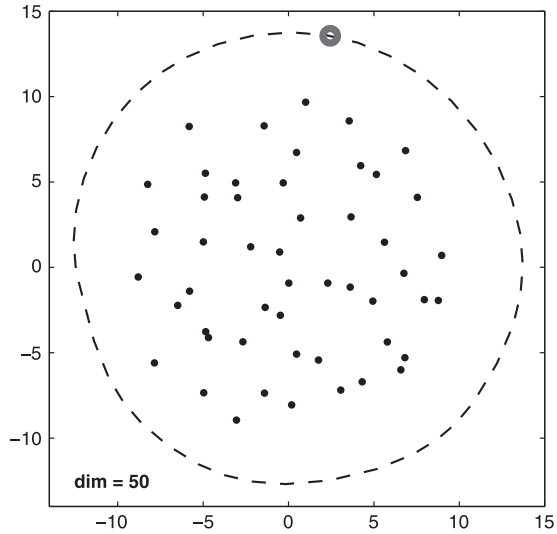

FIG. 2. The field of values of the matrix for which the pseudospectra is illustrated in Fig 1. A circle marks the point where the numerical radius is attained.

The lower bound together with the power inequality $r(A^k) \leqslant r(A)^k$ (Pearcy, 1966) implies an upper bound on the kth power of A

$$\|A^k\| \leqslant 2r(A)^k. \tag{1.7}$$

According to (1.7) one can gain insight about how large the norm of *any* power of A can be by computing $r(A)$. Moreover (1.6) implies that the upper bound $2r(A)^k$ on $\|A^k\|$ is typically tighter than $\|A\|^k$. Actually, $r(A)$ captures the norm of A as well as the asymptotic behaviour of the system (1.1). Therefore, it is a desirable measure for the analysis of the classical iterative systems for which the error can be represented by the recurrence (1.1). The analysis of the classical iterative methods using the field of values and the numerical radius has been studied by Axelsson *et al.* (1994) and Eiermann (1993).

For the pseudospectral radius it can be deduced from the Kreiss matrix theorem (Kreiss, 1962; Trefethen & Embree, 2005; Wegert & Trefethen, 1994) that

$$\sup_{\epsilon>0} \frac{\rho_\epsilon(A) - 1}{\epsilon} \leqslant \sup_k \|A^k\| \leqslant \mathrm{e}\, n \sup_{\epsilon>0} \frac{\rho_\epsilon(A) - 1}{\epsilon}. \tag{1.8}$$

In (1.8) the supremum of the norms of the matrix powers is bounded above and below in terms of the ϵ-pseudospectral radius. The lower bound is especially useful as an indicator of how large the norms of the matrix powers can grow. For instance, consider the 'Grcar' matrix (Trefethen & Embree, 2005), a 100×100 Toeplitz matrix with -1 on the first subdiagonal, 1 on the diagonal and on the first, second and third superdiagonals and all other entries 0. When we multiply this matrix by 0.4, we obtain a stable matrix A with $\rho(A) \approx 0.9052$. However, this matrix is very close to being unstable as is demonstrated by the fact that $\rho_\epsilon(A) \approx 1.0321$ for $\epsilon = 10^{-8}$. For this value of ϵ, the ratio $\frac{\rho_\epsilon(A)-1}{\epsilon}$ that lower bounds the left-hand side of (1.8) is approximately 3.2×10^6. For comparison purposes, a plot of the matrix power norms $\|A^k\|$ as a function of k is shown in Fig. 3. The norms decay to zero as they must since A is stable. Nonetheless, they first grow to the order of 10^7, which is approximately revealed by the

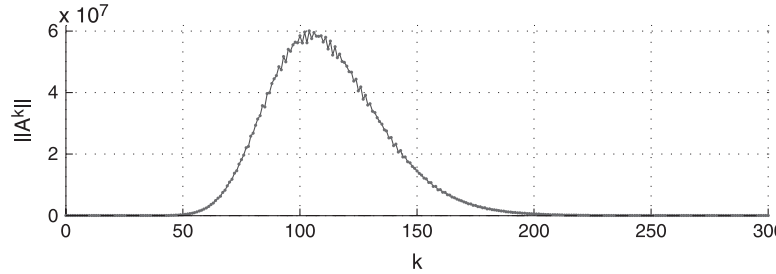

FIG. 3. The norms of the matrix powers for the 100 × 100 Grcar matrix multiplied by 0.4. The vertical axis corresponds to $\|A^k\|$ for the k value in the horizontal axis.

lower bound in (1.8) using $\epsilon = 10^{-8}$. See Trefethen & Embree (2005), Chapters 14 and 16 for extensive further discussion along these lines.

In this paper we develop algorithms to compute the robust stability measures $r(A)$ and $\rho_\epsilon(A)$. The algorithm we present for the latter is the first algorithm to compute the pseudospectral radius with high precision. For the computation of the numerical radius, the most recent method was suggested by He & Watson (1997). As we discuss in Section 3, the numerical radius can be reduced to an eigenvalue optimization problem with global maximum equal to the numerical radius. The method introduced in He & Watson (1997) is based on finding a local maximum of this eigenvalue optimization problem and verifying whether the local maximum is actually the global maximum by solving a generalized eigenvalue problem. The simple iteration introduced to locate a local maximum is not guaranteed to converge. Here we describe an algorithm that generates estimates converging to the numerical radius in exact arithmetic. The local convergence rate is usually quadratic. The algorithm is analogous to the Boyd–Balakrishnan algorithm for the $\mathbf{H}_\infty$ norm (Boyd & Balakrishnan, 1990) and depends on the solution of the generalized eigenvalue problems used for checking whether a local maximum is the global maximum in He & Watson (1997).

We put emphasis on the computation of the pseudospectral radius, the subject of Section 2. Before presenting a locally quadratically convergent algorithm for the ϵ-pseudospectral radius in Sections 2.2 and 2.3, we discuss its variational properties. The convergence analysis of the algorithm is simi- lar to that in Burke *et al.* (2003b), so we briefly justify our claims about its convergence properties in Section 2.4. The boundary of the pseudospectrum of a matrix may contain arcs of circles which may potentially cause numerical trouble for the pseudospectral radius algorithm. We investigate this phenomenon in Section 2.5. In Section 2.6 we specify a version of the algorithm in floating point arithmetic which is expected to produce accurate results as long as the pseudospectral radius problem is well conditioned. In Section 3 we describe the algorithm for the numerical radius without any convergence and accuracy analysis. We believe that the convergence analysis in Boyd & Balakrishnan (1990) and the accuracy analysis in this paper can be extended to this algorithm. The algorithm for the ϵ-pseudospectral radius is included in *EigTool*, Wright's state-of-the-art software for computing pseudospectra (Wright, 2002).

Notation. The smallest singular value of A is denoted by $\sigma_{\min}(A)$. The symbols $\bar{\lambda}$ and A^* denote the complex conjugate of the scalar λ and complex conjugate transpose of the matrix A, respectively. $\Lambda(A)$ and $\Lambda(A, B)$ are the set of eigenvalues of A and the set of generalized eigenvalues of the pencil $A - \lambda B$, respectively. When $M = M^*$, the largest eigenvalue of the matrix M is $\lambda_{\max}(M)$. When describing algorithms, we use the notation $x \bmod 2\pi$ which refers to the real number in the interval $[0, 2\pi)$ such that $x = k2\pi + x \bmod 2\pi$ for some integer k.

2. Computation of the pseudospectral radius

In this section we will utilize the well known fact (Trefethen & Embree, 2005) that (1.3) is equivalent to

$$\Lambda_\epsilon(A) = \{z \in \mathbb{C}: \sigma_{\min}(A - zI) \leqslant \epsilon\}. \tag{2.1}$$

We will also refer to the strict ϵ-pseudospectrum,

$$\Lambda'_\epsilon(A) = \{z \in \mathbb{C}: \sigma_{\min}(A - zI) < \epsilon\}. \tag{2.2}$$

For the convergence of continuous-time systems the analogous quantity to the pseudospectral radius is the pseudospectral abscissa, the maximum of the real parts of the points in the pseudospectrum. In Burke *et al.* (2003b), a quadratically convergent algorithm to compute the pseudospectral abscissa is given, so a first thought to compute the pseudospectral radius of A might be to reduce the problem to the computation of the pseudospectral abscissa of a related matrix. Given a complex number $r\,\mathrm{e}^{\mathrm{i}\theta}$ in the pseudospectrum, by taking the logarithm, we obtain $\ln r + \mathrm{i}\theta$. Denoting the set that is obtained by taking the logarithm of every point in $\Lambda_\epsilon(A)$ by $\ln(\Lambda_\epsilon(A))$, we conclude that the real part of the rightmost point in $\ln(\Lambda_\epsilon(A))$ is equal to the logarithm of the pseudospectral radius of A. However, there may not be any matrix with the pseudospectrum $\ln(\Lambda_\epsilon(A))$. For instance, as we discuss later in this section, there are matrices for which the boundary of the ϵ-pseudospectrum contains an arc of a circle centred at the origin. For such a matrix A, a line parallel to the imaginary axis intersects the boundary of the set $\ln(\Lambda_\epsilon(A))$ at infinitely many points, but in Burke *et al.* (2003b), it is shown that vertical cross sections of the ϵ-pseudospectrum of a matrix have only finitely many boundary points. Therefore, we derive an algorithm tailored to the pseudospectral radius, following the ideas in Burke *et al.* (2003b). Before presenting the algorithm, we discuss the variational properties of $\rho_\epsilon(A)$ that will play a crucial role in our analysis of the algorithm later.

2.1 *Variational properties of the ϵ-pseudospectral radius*

In this section we are interested in how the pseudospectral radius $\rho_\epsilon(X)$ varies with respect to ϵ and X. Thus, we view the pseudospectral radius as a mapping from $\mathbb{R}_+ \times \mathbb{C}^{n\times n}$ to $\mathbb{R}_+$. The most basic result we are looking for is the continuity of the pseudospectral radius with respect to ϵ and X. For this purpose we notice that the pseudospectral radius is the robust regularization of the spectral radius in the sense of Lewis (2002), i.e.

$$\rho_\epsilon(X) = \sup_Y\{\rho(Y): \|Y - X\| \leqslant \epsilon\}. \tag{2.3}$$

Since the spectral radius is continuous in matrix space, the continuity of $\rho_\epsilon(X)$ in matrix space immediately follows from Proposition 3.5 in Lewis (2002). In fact, joint continuity with respect to X and ϵ can also be shown by proving upper and lower semicontinuity separately (A. S. Lewis, personal communication).

THEOREM 2.1 (A. S. LEWIS) The function $\rho_\epsilon(X)$ is jointly continuous with respect to ϵ and X everywhere.

The next result states that the $(\epsilon + \beta)$-pseudospectral radius of X depends on the ϵ-pseudospectral radius of the matrices in the β neighbourhood of X.

THEOREM 2.2 Let β and ϵ be non-negative real numbers. Then the equality

$$\rho_{\epsilon+\beta}(X) = \sup_{\|X'-X\|\leqslant\beta} \rho_\epsilon(X')$$

holds.

Proof. By definition (2.3)

$$\begin{aligned}\rho_{\epsilon+\beta}(X) &= \sup_{Y}\{\rho(Y)\colon \|Y - X\| \leqslant \epsilon + \beta\}\\ &= \sup_{Y,X'}\{\rho(Y)\colon \|X' - X\| \leqslant \beta, \|Y - X'\| \leqslant \epsilon\}\\ &= \sup_{X'}\{\rho_\epsilon(X')\colon \|X' - X\| \leqslant \beta\}.\end{aligned}$$

Therefore, the result follows. □

Next we focus on the differentiability of $\rho_\epsilon(X)$. For this purpose let us introduce the function $p_{(\epsilon,X)}\colon \mathbb{R}_+ \times [0, 2\pi) \to \mathbb{R}$ for a given ϵ and X defined by

$$p_{(\epsilon,X)}(r,\theta) = \sigma_{\min}(X - r\,\mathrm{e}^{\mathrm{i}\theta} I) - \epsilon. \tag{2.4}$$

Note that $p_{(\epsilon,X)}(r,\theta)$ is less than or equal to 0 if and only if the complex number $r\,\mathrm{e}^{\mathrm{i}\theta}$ belongs to the ϵ-pseudospectrum of X. Well known properties of the minimum singular value function imply that $p_{(\epsilon,X)}(r,\theta)$ is a continuous function of r and θ. The theorem below specifies the conditions under which the function $p_{(\epsilon,X)}(r,\theta)$ is differentiable with respect to r and θ. Recall that a real-valued function defined on a real domain is called real-analytic at a given point if the function has a real convergent Taylor expansion at the given point.

THEOREM 2.3 Let $\epsilon \in \mathbb{R}_+$ and $X \in \mathbb{C}^{n\times n}$. If the minimum singular value of $X - r\,\mathrm{e}^{\mathrm{i}\theta} I$ is greater than 0 and has multiplicity one, then at (r,θ) the function $p_{(\epsilon,X)}(r',\theta')$ is real-analytic with respect to r' and θ' with derivatives

$$\nabla p_{(\epsilon,X)}(r,\theta) = (-\mathrm{Re}(\mathrm{e}^{\mathrm{i}\theta} u^* v), \mathrm{Im}(r\,\mathrm{e}^{\mathrm{i}\theta} u^* v)),$$

where u and v are any consistent pair of unit left and right singular vectors corresponding to $\sigma_{\min}(X - r\,\mathrm{e}^{\mathrm{i}\theta} I)$.

Proof. The function $\sigma_{\min}(X - r'\,\mathrm{e}^{\mathrm{i}\theta'} I)$ is real-analytic at (r,θ) provided $\sigma_{\min}(X - r\,\mathrm{e}^{\mathrm{i}\theta} I)$ is positive and has multiplicity one. This immediately follows from the fact that $X_2(r',\theta') = (X^* - r'\,\mathrm{e}^{-\mathrm{i}\theta'} I)(X - r'\,\mathrm{e}^{\mathrm{i}\theta'} I)$ is analytic with respect to r' and θ' and therefore $\sigma^2_{\min}(X - r'\,\mathrm{e}^{\mathrm{i}\theta'} I)$, the smallest eigenvalue of $X_2(r',\theta')$, is real-analytic whenever $\sigma^2_{\min}(X - r'\,\mathrm{e}^{\mathrm{i}\theta'} I)$ has multiplicity one. The derivatives can be derived by applying the chain rule to the result of Theorem 7.1 in Burke *et al.* (2003a). □

For a fixed ϵ and X, we call the constrained optimization problem

$$\sup_{p_{(\epsilon,X)}(r,\theta)\leqslant 0} r, \tag{2.5}$$

the *ϵ-pseudospectral radius problem at X*. By the definition of $p_{(\epsilon,X)}(r,\theta)$ and the definition of the ϵ-pseudospectral radius given in (2.1), we see that the value attained at a global maximizer of the ϵ-pseudospectral radius problem at X is equal to $\rho_\epsilon(X)$. Now we are ready to derive the derivatives of $\rho_\epsilon(X)$ with respect to ϵ and X.

THEOREM 2.4 Let a matrix $X_0 \in \mathbb{C}^{n\times n}$ and $\epsilon_0 \in \mathbb{R}_+$ be given. Suppose (r_0,θ_0) is a local maximizer of the ϵ_0-pseudospectral radius problem at X_0 and the multiplicity of $\sigma_{\min}(X_0 - r_0\,\mathrm{e}^{\mathrm{i}\theta_0} I)$ is one. Then the gradient of $p_{(\epsilon_0,X_0)}(r,\theta)$ at (r_0,θ_0) is a non-negative multiple of $(1, 0)$.

Moreover if the point (r_0,θ_0) is the unique global maximizer, $\nabla p_{(\epsilon_0,X_0)}(r_0,\theta_0)$ is non-zero and the Hessian of p with respect to r and θ, $\nabla^2 p_{(\epsilon_0,X_0)}(r_0,\theta_0)$, is non-singular, then at (ϵ_0, X_0) the function

$\rho_\epsilon(X)$ is differentiable with respect to ϵ and X with derivatives

$$\frac{\mathrm{d}\rho_{\epsilon_0}(X_0)}{\mathrm{d}\epsilon} = \frac{-1}{\mathrm{Re}(\mathrm{e}^{\mathrm{i}\theta} u^* v)}, \quad \nabla_X \rho_{\epsilon_0}(X_0) = \frac{u v^*}{\mathrm{Re}(\mathrm{e}^{\mathrm{i}\theta} u^* v)},$$

where u and v are any consistent pair of unit left and right singular vectors corresponding to $\sigma_{\min}(X_0 - r_0\,\mathrm{e}^{\mathrm{i}\theta_0} I)$.

Proof. Since (r_0, θ_0) is a local maximizer of the ϵ_0-pseudospectral radius problem at X_0 and by Theorem 2.3 $p_{(\epsilon_0, X_0)}(r, \theta)$ is differentiable with respect to r and θ at this maximizer, provided its gradient is non-zero, standard first-order necessary conditions must be satisfied. Thus, either the gradient of $p_{(\epsilon_0, X_0)}$ at (r_0, θ_0) is 0 or there exists a positive μ such that

$$(1, 0) - \mu \nabla p_{(\epsilon_0, X_0)}(r_0, \theta_0) = 0.$$

In either case the gradient is a non-negative multiple of $(1, 0)$ as desired. From Theorem 2.3, we know that

$$\nabla p_{(\epsilon_0, X_0)}(r_0, \theta_0) = (-\mathrm{Re}(\mathrm{e}^{\mathrm{i}\theta_0} u^* v), \mathrm{Im}(r_0 \mathrm{e}^{\mathrm{i}\theta_0} u^* v)),$$

so when $u^* v \neq 0$, we have $\mu = \frac{-1}{\mathrm{Re}(\mathrm{e}^{\mathrm{i}\theta_0} u^* v)}$.

When (r_0, θ_0) is the unique global maximizer with non-zero gradient and non-singular Hessian, we deduce from a standard sensitivity result such as Theorem 5.53 in Bonnans & Shapiro (2000) that $\frac{\mathrm{d}\rho_{\epsilon_0}(X_0)}{\mathrm{d}\epsilon} = -\mu \frac{\mathrm{d}p_{(\epsilon_0, X_0)}(r_0, \theta_0)}{\mathrm{d}\epsilon} = \mu$ and $\nabla_X \rho_{\epsilon_0}(X_0) = -\mu \nabla_X p_{(\epsilon_0, X_0)}(r_0, \theta_0) = -\mu u v^*$ hold (since $\nabla_X p_{(\epsilon_0, X_0)}(r_0, \theta_0) = u v^*$; see Theorem 7.1 in Burke *et al.* (2003a)). □

2.2 *Radial and circular searches*

The algorithm depends on the steps that we call circular and radial searches. Figure 4 illustrates a radial and a circular search for a variant of a 3×3 example given by Demmel (1987) and for $\epsilon = 10^{-3.18}$. This matrix is an upper triangular Toeplitz matrix with the entry $d_{j,k}$, $k > j$, equal to $-10^{2(k-j)}$ and the entries on the diagonal equal to $0.1 + 0.01\mathrm{i}$. For the rest of this section, let us fix $\epsilon \in \mathbb{R}_+$ and the matrix $A \in \mathbb{C}^{n \times n}$ for which we are computing the pseudospectral radius. We drop the subscripts of the function $p_{(\epsilon, A)}(r, \theta)$ for convenience.

The aim of a radial search is to find the point on the boundary of the ϵ-pseudospectrum with the largest modulus in a given direction. More formally, given $\theta \in [0, 2\pi)$ such that there exists a positive real number r' satisfying $p(r', \theta) = 0$, we want to calculate

$$\eta_\epsilon(\theta) = \max\{r \in \mathbb{R}_+ \colon p(r, \theta) = 0\}. \tag{2.6}$$

We will state a theorem which suggests how we can compute the r values such that $p(r, \theta) = 0$ holds for a fixed $\theta \in [0, 2\pi)$.

THEOREM 2.5 Let $r, \epsilon \in \mathbb{R}_+$ and $\theta \in [0, 2\pi)$. The matrix $A - r\,\mathrm{e}^{\mathrm{i}\theta} I$ has ϵ as one of its singular values if and only if the matrix

$$K(\theta, \epsilon) = \begin{bmatrix} \mathrm{i}\,\mathrm{e}^{\mathrm{i}\theta} A^* & \epsilon I \\ -\epsilon I & \mathrm{i}\,\mathrm{e}^{-\mathrm{i}\theta} A \end{bmatrix} \tag{2.7}$$

has the pure imaginary eigenvalue $\mathrm{i}r$.

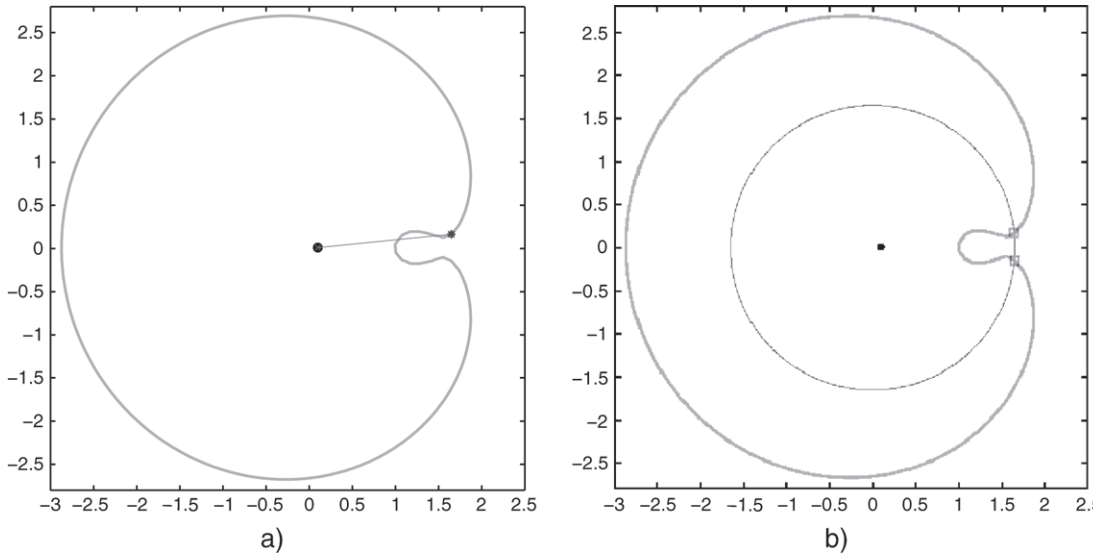

FIG. 4. The boundary of the ϵ-pseudospectrum for an example due to Demmel. (a) The radial search finds the point with the maximum modulus on the pseudospectrum boundary in a given search direction. (b) The circular search determines the intersection points of the ϵ-pseudospectrum boundary with a circle of given radius.

Proof. The matrices $A - r\,\mathrm{e}^{\mathrm{i}\theta} I$ and $\mathrm{i}A\,\mathrm{e}^{-\mathrm{i}\theta} - \mathrm{i}r I$ have the same set of singular values. It follows from Byers (1988, Theorem 1) and Burke *et al.* (2003a, Lemma 5.3) that the matrix $\mathrm{i}A\,\mathrm{e}^{-\mathrm{i}\theta} - \mathrm{i}r I$ has the singular value ϵ if and only if the imaginary number $\mathrm{i}r$ is an eigenvalue of the matrix in (2.7). □

We note that the matrix $K(\theta, \epsilon)$ is Hamiltonian, i.e. $JK(\theta, \epsilon)$ is Hermitian, where

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}, \tag{2.8}$$

with $n \times n$ blocks. Using definition (2.6) and Theorem 2.5, $\eta_\epsilon(\theta)\mathrm{i}$ is an imaginary eigenvalue of $K(\theta, \epsilon)$. According to the next corollary $\eta_\epsilon(\theta)\mathrm{i}$ is actually the imaginary eigenvalue with the largest imaginary part.

COROLLARY 2.6 (RADIAL SEARCH) Given a number $\theta \in [0, 2\pi)$ with $p(r', \theta) = 0$ for some r', the quantity $\eta_\epsilon(\theta)$ defined in (2.6) is the largest of the imaginary parts of the pure imaginary eigenvalues of $K(\theta, \epsilon)$.

Proof. Since there exists r' such that $p(r', \theta) = 0$, Theorem 2.5 implies that the matrix $K(\theta, \epsilon)$ has an imaginary eigenvalue. Let $r_\epsilon(\theta)\mathrm{i}$ be the imaginary eigenvalue of the matrix $K(\theta, \epsilon)$ with greatest imaginary part. Using definition (2.6) and Theorem 2.5, $\eta_\epsilon(\theta)\mathrm{i} \in \Lambda(K(\theta, \epsilon))$, i.e. $r_\epsilon(\theta) \geqslant \eta_\epsilon(\theta)$. Now suppose $r_\epsilon(\theta)$ is strictly greater than $\eta_\epsilon(\theta)$. Again from Theorem 2.5, we deduce that $A - r_\epsilon(\theta)\,\mathrm{e}^{\mathrm{i}\theta} I$ has a singular value ϵ (not necessarily the smallest one), so $p(r_\epsilon(\theta), \theta) \leqslant 0$. Since p is a continuous function of r and $p(r, \theta)$ approaches ∞ as r goes to ∞, from the intermediate value theorem we conclude that for some $\hat{r} \geqslant r_\epsilon(\theta) > \eta_\epsilon(\theta)$, $p(\hat{r}, \theta) = 0$. But this contradicts the definition of $\eta_\epsilon(\theta)$ in (2.6). Therefore, $\eta_\epsilon(\theta) = r_\epsilon(\theta)$ must hold. □

In a circular search we identify the set of points on the boundary of the pseudospectrum with a given modulus. In other words, given a positive real number r, we need to determine those θ values in the

interval $[0, 2\pi)$ for which $p(r, \theta) = 0$ is satisfied. A result from Byers (1988) implies that $A - \mathrm{e}^{\mathrm{i}\theta} I$ has ϵ as one of its singular values if and only if the pencil $P(1, \epsilon) - \lambda Q(1, \epsilon)$ has the generalized eigenvalue $\mathrm{e}^{\mathrm{i}\theta}$, where

$$P(r, \epsilon) = \begin{bmatrix} -\epsilon I & A \\ r I & 0 \end{bmatrix}, \quad Q(r, \epsilon) = \begin{bmatrix} 0 & r I \\ A^* & -\epsilon I \end{bmatrix}. \tag{2.9}$$

The pencil $P(r, \epsilon)^* - \lambda Q(r, \epsilon)^*$ is symplectic, i.e. $P(r, \epsilon)^* J P(r, \epsilon) = Q(r, \epsilon)^* J Q(r, \epsilon)$ for the matrix J defined in (2.8). Apart from the symplectic structure of the pencil $P(r, \epsilon) - \lambda Q(r, \epsilon)$, we note that $D(\theta)(P(r, \epsilon) - \mathrm{e}^{\mathrm{i}\theta} Q(r, \epsilon))$ is Hermitian for all θ, where

$$D(\theta) = \begin{bmatrix} I & 0 \\ 0 & -\mathrm{e}^{-\mathrm{i}\theta} I \end{bmatrix}. \tag{2.10}$$

The error analysis in Section 2.6 exploits this structure.

We present a generalized version of Byers' result, establishing a relation between the singular values of $A - r\,\mathrm{e}^{\mathrm{i}\theta} I$ and the eigenvalues of the pencil $P(r, \epsilon) - \lambda Q(r, \epsilon)$. We recall that a $2n \times 2n$ pencil $X - \lambda Y$ is said to be singular if $\det(X - \lambda Y) = 0$ for all $\lambda \in \mathbb{C}$; otherwise it is said to be regular in which case it has at most $2n$ finite eigenvalues.

THEOREM 2.7 (CIRCULAR SEARCH) The matrix $A - r\,\mathrm{e}^{\mathrm{i}\theta} I$ has ϵ as one of its singular values if and only if the pencil $P(r, \epsilon) - \lambda Q(r, \epsilon)$ has the generalized eigenvalue $\mathrm{e}^{\mathrm{i}\theta}$ or the pencil $P(r, \epsilon) - \lambda Q(r, \epsilon)$ is singular.

Proof. The matrix $A - r\,\mathrm{e}^{\mathrm{i}\theta} I$ has the singular value ϵ if and only if

$$\begin{bmatrix} 0 & A - r\,\mathrm{e}^{\mathrm{i}\theta} I \\ A^* - r\,\mathrm{e}^{-\mathrm{i}\theta} I & 0 \end{bmatrix}$$

has ϵ as one of its eigenvalues. But this holds if and only if

$$\det\begin{bmatrix} -\epsilon I & A - r\,\mathrm{e}^{\mathrm{i}\theta} I \\ A^* - r\,\mathrm{e}^{-\mathrm{i}\theta} I & -\epsilon I \end{bmatrix} = 0$$

or equivalently, multiplying the matrix above by $D^*(\theta)$ on the left

$$\det\begin{bmatrix} -\epsilon I & A - r\,\mathrm{e}^{\mathrm{i}\theta} I \\ -\mathrm{e}^{\mathrm{i}\theta} A^* + r I & \epsilon\,\mathrm{e}^{\mathrm{i}\theta} I \end{bmatrix} = 0.$$

By rearranging the matrix above, we see that $\det(P(r, \epsilon) - \mathrm{e}^{\mathrm{i}\theta} Q(r, \epsilon)) = 0$. □

Unlike a radial search, in a circular search we wish to determine all of the zeros of $p(r, \cdot)$. Hence, as long as $P(r, \epsilon) - \lambda Q(r, \epsilon)$ is regular, to find the points on the ϵ-pseudospectrum boundary with modulus r we need to check whether $A - r\,\mathrm{e}^{\mathrm{i}\theta} I$ has ϵ as its minimum singular value for each $\theta \in [0, 2\pi)$ such that $\mathrm{e}^{\mathrm{i}\theta} \in \Lambda(P(r, \epsilon), Q(r, \epsilon))$.

2.3 *The algorithm*

We now combine radial and circular searches to obtain an algorithm for the ϵ-pseudospectral radius. For now, we assume that the pencil we use for circular searches is regular for all values of r. The issue of singular pencils is the theme of Section 2.5. In particular, we explain how the algorithm below can be modified for singular pencils.

The algorithm is based on the Boyd–Balakrishnan algorithm (Boyd & Balakrishnan, 1990) and the criss-cross method for the pseudospectral abscissa introduced by Burke *et al.* (2003b). It keeps an

estimate of the pseudospectral radius and a set of open 'intervals', $I_1^j, I_2^j, \ldots, I_{m_j}^j$. Actually, all these are intervals $(\gamma_k^j, \zeta_k^j) \subset [0, 2\pi)$ with the possible exception of $I_{m_j}^j$ which may 'wrap around the circle', i.e. $I_{m_j}^j = (\gamma_{m_j}^j, 2\pi) \cup [0, \zeta_{m_j}^j)$ with $\gamma_{m_j}^j > \zeta_{m_j}^j$. Let the real number η^j be the estimate of the pseudospectral radius at the jth iteration and let $\theta \in [0, 2\pi)$. Then for $j > 1$ the point $\eta^j \, \mathrm{e}^{\mathrm{i}\theta}$ lies inside the strict pseudospectrum if and only if the angle θ is contained in one of $I_1^j, I_2^j, \ldots, I_{m_j}^j$.

At each iteration, the algorithm applies a radial search in the direction of the midpoint of each interval. The estimate of the pseudospectral radius is refined to the maximum of the modulus values returned by the radial searches. The open intervals are updated by the application of a circular search. New open intervals contain the angles of the points lying inside the strict pseudospectrum and on the circle with radius equal to the new estimate of the pseudospectral radius. Initially, we start with a radial search in the direction of the angle of an arbitrary eigenvalue whose modulus is equal to the spectral radius.

ALGORITHM 2.8

1. Let θ_ρ be the angle of an eigenvalue with modulus $\rho(A)$, set $j = 0$ and $\Phi^0 = [\theta_\rho]$.
2. Radial searches: Perform a radial search for each midpoint $\Phi_k^j \in \Phi^j$. Compute

$$\eta^{j+1} = \max\{\eta_\epsilon(\Phi_k^j) \colon \Phi_k^j \in \Phi^j\}, \tag{2.11}$$

   where η_ϵ is defined in (2.6).
3. Circular search: Perform a circular search to find the intersection points of the circle with radius η^{j+1} and the ϵ-pseudospectrum boundary. Using these intersection points determine the open intervals $I_1^{j+1}, I_2^{j+1}, \ldots, I_{m^{j+1}}^{j+1}$ in which $p(\eta^{j+1}, \cdot)$ is negative. Compute the new set of midpoints

$$\Phi^{j+1} = \{\Phi_1^{j+1}, \Phi_2^{j+1}, \ldots, \Phi_{m^{j+1}}^{j+1}\},$$

   where Φ_k^{j+1} is the midpoint of the interval I_k^{j+1},

$$\Phi_k^{j+1} = \begin{cases} \frac{\gamma_k^{j+1} + \zeta_k^{j+1}}{2}, & \text{if } \gamma_k^{j+1} < \zeta_k^{j+1}, \\ \frac{\gamma_k^{j+1} + \zeta_k^{j+1} + 2\pi}{2} \bmod 2\pi, & \text{otherwise.} \end{cases}$$

4. Increment j by one, go to step 2.

In Fig 5, the first two iterations of a sample run of the algorithm are shown. The initial radial search is followed by a circular search which detects four intersection points. Next we perform two radial searches in the directions of the midpoints of two intervals in which $p(\eta^1, \cdot)$ is negative. The maximum of the values returned by the radial searches is our next estimate η^2 for the ϵ-pseudospectral radius. For the specific example, the input matrix is real, so the values returned by the radial searches are equal. We continue with a circular search as before.

It is possible to obtain a slight improvement in Algorithm 2.8 by changing the radial search to return the largest r in absolute value such that $p(r, \theta) = 0$. Corollary 2.6 can be extended to show that the modulus of the pure imaginary eigenvalue of $K(\theta, \epsilon)$ with the largest imaginary part in absolute value is the largest zero of $p(\cdot, \theta)$ in absolute value. This version of the radial search may occasionally provide

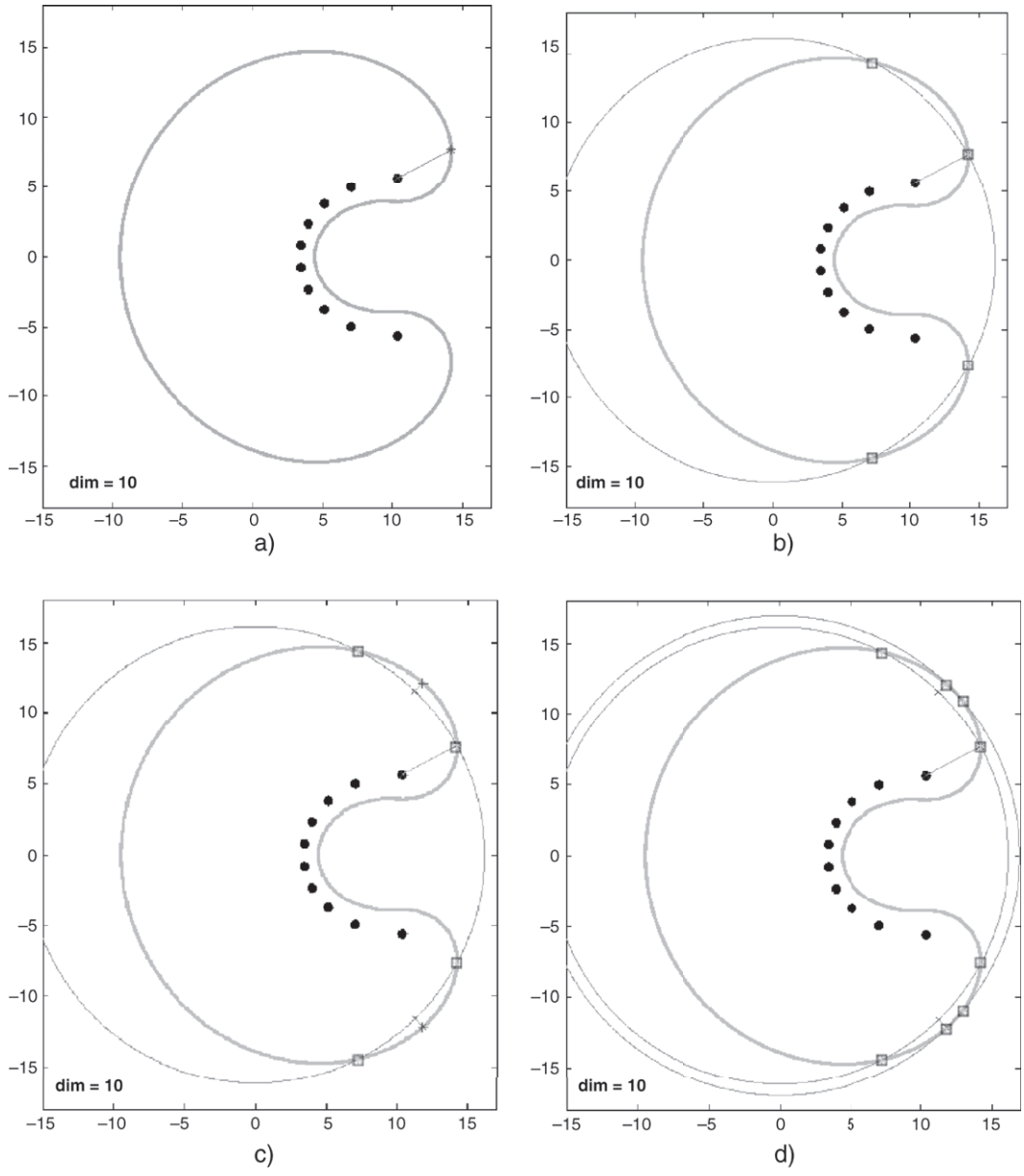

FIG. 5. First two iterations of the algorithm on a shifted companion matrix.

a better initial estimate, however, for the later iterations the gain is likely to be insignificant. To keep the description and analysis simple we use definition (2.6).

In Algorithm 2.8 one point that is left unspecified is how the intervals $I_1^j, I_2^j, \ldots, I_{m_j}^j$ can be determined from the intersection points returned by a circular search. One trivial and robust way is to sort the

intersection points and compute $\sigma_{\min}(A - \eta^j \, \mathrm{e}^{\mathrm{i}\theta} I)$ at the midpoint θ of each adjacent pair. The adjacent pair constitutes an interval in which $p(\eta^j, \cdot) < 0$ is satisfied if and only if $\sigma_{\min}(A - \eta^j \, \mathrm{e}^{\mathrm{i}\theta} I) < \epsilon$. Another possibility is to classify the intersection points as crossing or non-crossing zeros. We call the intersection point θ' a crossing zero of $p(r, \cdot)$ if $p(r, \cdot)$ has opposite sign on $(\theta' - \alpha, \theta']$ and $[\theta', \theta' + \alpha')$ for sufficiently small positive α. Otherwise the intersection point is called a non-crossing zero of $p(r, \cdot)$. We can distinguish the crossing zeros from non-crossing zeros using the theorem below under the assumption that $\sigma_{\min}(A - r \, \mathrm{e}^{\mathrm{i}\theta} I)$ is of multiplicity one for each intersection point $r \, \mathrm{e}^{\mathrm{i}\theta}$.

THEOREM 2.9 (CROSSING VERSUS NON-CROSSING ZEROS DURING THE CIRCULAR SEARCHES) Let $r \in \mathbb{R}_+$ and $\mathrm{e}^{\mathrm{i}\theta_0}$ be an eigenvalue of the pencil $P(r, \epsilon) - \lambda Q(r, \epsilon)$. Moreover suppose that $\sigma_{\min}(A - r \, \mathrm{e}^{\mathrm{i}\theta_0} I)$ is simple and equal to ϵ. Then θ_0 is a crossing zero of $p(r, \cdot)$ if and only if the algebraic multiplicity of the eigenvalue $\mathrm{e}^{\mathrm{i}\theta_0}$ is odd.

Proof. By the definitions of $P(r, \epsilon)$ and $Q(r, \epsilon)$ (see (2.9))

$$P(r, \epsilon) - \lambda Q(r, \epsilon) = \det \begin{bmatrix} -\epsilon I & A - \lambda r I \\ r I - \lambda A^* & \lambda \epsilon I \end{bmatrix}.$$

We define the function $q \colon \mathbb{C} \to \mathbb{C}$ as the determinant of this matrix with the bottom block multiplied by $-\bar{\lambda}$,

$$q(\lambda) = (-1)^n \bar{\lambda}^n \det(P(r, \epsilon) - \lambda Q(r, \epsilon)) = \det \begin{bmatrix} -\epsilon I & A - \lambda r I \\ |\lambda|^2 A^* - \bar{\lambda} r I & -|\lambda|^2 \epsilon I \end{bmatrix}. \tag{2.12}$$

Define a function $g \colon \mathbb{R} \to \mathbb{C}$ by $g(\theta) = q(\mathrm{e}^{\mathrm{i}\theta})$. Now if the multiplicity of $\mathrm{e}^{\mathrm{i}\theta_0}$ as the eigenvalue of the pencil $P(r, \epsilon) - \lambda Q(r, \epsilon)$ is m, we have

$$g(\theta) = q(\mathrm{e}^{\mathrm{i}\theta}) = \beta(\theta)(\mathrm{e}^{\mathrm{i}\theta} - \mathrm{e}^{\mathrm{i}\theta_0})^m, \tag{2.13}$$

where $\beta \colon \mathbb{R} \to \mathbb{C}$ is a continuous function with $\beta(\theta_0) \neq 0$. Furthermore, when we make the substitution $\lambda = \mathrm{e}^{\mathrm{i}\theta}$ in the right-hand side of (2.12), we see that the eigenvalues of the resulting matrix are $\pm\sigma_j(A - r \, \mathrm{e}^{\mathrm{i}\theta} I) - \epsilon$, i.e. plus and minus the singular values of $A - r \, \mathrm{e}^{\mathrm{i}\theta} I$ decremented by ϵ. Therefore,

$$g(\theta) = (-1)^n \prod_{j=1}^{n} (\sigma_j(A - r \, \mathrm{e}^{\mathrm{i}\theta} I) - \epsilon)(\sigma_j(A - r \, \mathrm{e}^{\mathrm{i}\theta} I) + \epsilon) \tag{2.14}$$

implying $g(\theta)$ is real valued for all θ.

Now for real small α, we deduce from the equality

$$\mathrm{e}^{\mathrm{i}(\theta_0 + \alpha)} - \mathrm{e}^{\mathrm{i}\theta_0} = \mathrm{e}^{\mathrm{i}\theta_0}(\mathrm{e}^{\mathrm{i}\alpha} - 1) = \mathrm{e}^{\mathrm{i}\theta_0}\mathrm{i}\alpha + \mathrm{O}(\alpha^2) = \alpha(\mathrm{e}^{\mathrm{i}(\theta_0 + \pi/2)} + \mathrm{O}(\alpha))$$

and from (2.13) that

$$g(\theta_0 + \alpha) = \alpha^m(\beta(\theta_0 + \alpha) \, \mathrm{e}^{m\mathrm{i}(\theta_0 + \pi/2)} + \mathrm{O}(\alpha)) \equiv \alpha^m f(\theta_0, \alpha)$$

holds, where $f \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous. Notice that because of the continuity of f and the fact that $f(\theta_0, 0) = \beta(\theta_0) \, \mathrm{e}^{m\mathrm{i}(\theta_0 + \pi/2)}$ is a non-zero real number, $f(\theta_0, \alpha)$ and $f(\theta_0, -\alpha)$ are non-zero with the same sign. Therefore, for all small α, $g(\theta + \alpha) = \alpha^m f(\theta_0, \alpha)$ and $g(\theta - \alpha) = (-\alpha)^m f(\theta_0, -\alpha)$ have different signs if and only if m is odd. But according to (2.14) the sign of $g(\theta)$ changes around θ_0 if and only if the sign of $p(r, \theta)$ changes. □

Theorem 2.9 allows us in principle to classify in which intervals $p(\eta^j, \cdot)$ is negative by evaluating $\sigma_{\min}(A - \eta^j \, \mathrm{e}^{\mathrm{i}\theta} I)$ only at the midpoint of one pair of intersection points computed in step 3, provided the assumption that $\sigma_{\min}(A - r \, \mathrm{e}^{\mathrm{i}\theta} I)$ is simple at the intersection points is valid. In practice, however, evaluation of $\sigma_{\min}(A - \eta^j \, \mathrm{e}^{\mathrm{i}\theta} I)$ at every midpoint seems a simpler and more robust way to determine in which intervals $p(\eta^j, \cdot)$ is negative.

2.4 *Convergence analysis of Algorithm 2.8*

We claim that the sequence of iterates $\{\eta^j\}$ generated by Algorithm 2.8 converges to the ϵ-pseudospectral radius of A. Recall that we assume the pencil for the circular searches is regular which implies that there are at most $2n$ intersection points of the circle of radius r and the ϵ-pseudospectrum boundary. The convergence proof is analogous to that of the crisscross method to compute the pseudospectral abscissa (Burke *et al.*, 2003b, Theorem 3.2). Therefore, we shall just give an outline of the proof.

First note that on a circle centered at the origin and with radius strictly between the spectral radius and the ϵ-pseudospectral radius there are points lying in the strict ϵ-pseudospectrum as shown by the following argument. Given a point z on the boundary of the ϵ-pseudospectrum, according to definition (1.3) $z \in \Lambda(A + E)$ for some E with norm ϵ. But the eigenvalues of $A + tE$ are continuous functions of $t \in [0, 1]$. Therefore, there must be a continuous path from each point on the ϵ-pseudospectrum boundary to an eigenvalue of A lying entirely in the strict ϵ-pseudospectrum except the end point on the boundary.

If at some iteration j the ϵ-pseudospectral radius estimate η^j is equal to the ϵ-pseudospectral radius, there is nothing to prove. Thus, suppose that none of the estimates is equal to the ϵ-pseudospectral radius. In this case the estimates $\{\eta^j\}$ are monotonically increasing, bounded above by the ϵ-pseudospectral radius and bounded below by the spectral radius. This can be easily shown by induction considering the update rule (2.11) and the definition of $\eta_\epsilon(\theta)$ in (2.6).

Since the estimates are in increasing order bounded above by the ϵ-pseudospectral radius, they must converge to a real number η^∞ less than or equal to the ϵ-pseudospectral radius. Suppose η^∞ is strictly less than the ϵ-pseudospectral radius. There must be open intervals such that the function $p(\eta^\infty, \theta)$ is non-positive. Otherwise we obtain a contradiction with the result stating that for all r between the spectral radius and the pseudospectral radius there are points lying inside the ϵ-pseudospectrum and on the circle centered at the origin with radius r. But from the existence of the open intervals in which the inequality $p(\eta^\infty, \theta) \leqslant 0$ is satisfied, it is possible to deduce $p(\eta^\infty, \Phi_k^j) \leqslant 0$ for sufficiently large j and for some k. Therefore, the inequality $\eta^{j+1} \geqslant \eta^\infty$ holds for sufficiently large j. This contradicts the fact that the iterates are monotonically increasing. Therefore, the limit η^∞ must be equal to the ϵ-pseudospectral radius. Thus we have:

THEOREM 2.10 Suppose the pencil $P(r, \epsilon) - \lambda Q(r, \epsilon)$ is regular for all positive r. Then the sequence $\{\eta^j\}$ generated by Algorithm 2.8 converges to $\rho_\epsilon(A)$.

Just like the crisscross algorithm for the pseudospectral abscissa, we expect Algorithm 2.8 to converge to the pseudospectral radius quadratically under the same regularity assumption stated in Burke *et al.* (2003b), namely, that the global maximizers of the ϵ-pseudospectral radius problem (2.5) are regular. In Burke *et al.* (2003b) a point in the complex plane (x, y) is called regular if the multiplicity of the minimum singular value of $A - (x + \mathrm{i}y)I$ is one and the pair of left and right singular vectors corresponding to this minimum singular value are not orthogonal to each other. To show the quadratic convergence, the approach in Burke *et al.* (2003b, Section 4 and Section 5) can be followed. The crucial point that is worth noting here is that by Theorem 2.3 the function $p(r, \theta)$ is analytic whenever the

minimum singular value of $A - r\,\mathrm{e}^{\mathrm{i}\theta} I$ is positive and has multiplicity one. Additionally, by Theorem 2.4 around a regular local maximum of the pseudospectral radius problem, the gradient of p must be a positive multiple of $(1, 0)$. Suppose the point (r_0, θ_0) is a regular local maximum. Now an analogous argument to that of Theorem 4.1 and Corollary 4.5 in Burke *et al.* (2003b) applies to deduce the existence of a real-analytic function $f(\theta)$ near zero such that $p(r, \theta)$ and $r - r_0 + f(\theta - \theta_0)$ have the same signs for all r and θ sufficiently close to (r_0, θ_0). Moreover the function f satisfies the properties

$$f(0) = f'(0) = \cdots = f^{(2k-1)}(0) = 0, \quad f^{(2k)}(0) > 0, \tag{2.15}$$

for some $k \geqslant 1$. According to Section 5 in Burke *et al.* (2003b) since the pseudospectrum around a local maximum can be described by a function satisfying (2.15), Algorithm 2.8 converges quadratically to the global maximum which is the pseudospectral radius in our case.

As argued in Burke *et al.* (2003b), for generic matrices the multiplicity of $\sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta} I)$ is one at the maximizer (r, θ). If the multiplicity of the minimum singular value is greater than one at a maximizer, the quadratic convergence proof outlined above does not apply, although it may be possible to extend the proof to cover such cases.

2.5 *Singular pencils in the circular search*

We first consider the geometrical interpretation of the singularity of the pencil in a circular search. When the boundary of the ϵ-pseudospectrum of A contains an arc of the circle of radius r centered at the origin, we infer from Theorem 2.7 that the pencil $P(r, \epsilon) - \lambda Q(r, \epsilon)$ is singular. Notice that the reverse implication does not necessarily hold. For generic matrices the minimum singular value of $A - r\,\mathrm{e}^{\mathrm{i}\theta} I$ has multiplicity one for all θ (see Burke *et al.*, 2003b) and Theorem 2.11 tells us that there are actually only two possibilities when the pencil $P(r, \epsilon) - \lambda Q(r, \epsilon)$ is singular.

THEOREM 2.11 (SINGULAR PENCILS AND CIRCULAR PSEUDOSPECTRA) Given a positive real number r suppose that the pencil $P(r, \epsilon) - \lambda Q(r, \epsilon)$ is singular and that $\sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta} I)$ has multiplicity one for all $\theta \in [0, 2\pi)$. Then either

- the circle with radius r lies completely inside the strict ϵ-pseudospectrum or
- the ϵ-pseudospectrum boundary contains the circle of radius r.

Proof. By Theorem 2.7 the singularity of the pencil guarantees that given an arbitrary $\theta \in [0, 2\pi)$, the matrix $A - r\,\mathrm{e}^{\mathrm{i}\theta} I$ has ϵ as one of its singular values, so $p(r, \theta) \leqslant 0$. If for all θ, $p(r, \theta) < 0$ is satisfied, the first case holds. So assume there is a zero of $p(r, \cdot)$. By way of contradiction, suppose there exists $\tilde{\theta}$ such that $p(r, \tilde{\theta}) < 0$. Let $\hat{\theta}$ be the zero of $p(r, \cdot)$ closest to $\tilde{\theta}$. Without loss of generality assume $\hat{\theta}$ is greater than $\tilde{\theta}$. For all $\theta \in [\tilde{\theta}, \hat{\theta})$, $p(r, \theta) < 0$, so the smallest singular value of $A - r\,\mathrm{e}^{\mathrm{i}\theta} I$ is strictly less than ϵ, and hence, the second smallest singular value of $A - r\,\mathrm{e}^{\mathrm{i}\theta} I$ is less than or equal to ϵ. It follows by the continuity of singular values that the second smallest singular value of $A - r\,\mathrm{e}^{\mathrm{i}\hat{\theta}} I$ is less than or equal to ϵ. This contradicts the fact that $\sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\hat{\theta}} I)$ is equal to ϵ and has multiplicity one. Thus, $p(r, \theta) = 0$ for all θ, so the second case holds. □

Now returning to Algorithm 2.8, we note that for all j there is a zero of the function $p(\eta^j, \cdot)$ because of the way we update the estimates of the pseudospectral radius (2.11). Therefore, the circle of radius η^j cannot completely lie inside the strict pseudospectrum. In other words, for generic matrices the singularity of the pencil used by Algorithm 2.8 for the circular search implies that the ϵ-pseudospectrum boundary contains a circle.

In general, the presence of singular pencils is not desirable for Algorithm 2.8, because it is difficult to determine the singularity of a pencil. Thus, our strategy to handle singular pencils is to try to avoid them. This turns out to be surprisingly simple. The next result is a corollary of Theorem 2.7.

COROLLARY 2.12 (AVOIDING SINGULAR PENCILS) Let r be a positive real number such that $\sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta} I) > \epsilon$ for some θ. Then the pencil $P(r,\epsilon) - \lambda Q(r,\epsilon)$ is regular.

For all r greater than η^1, by (2.6) $\sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta_\rho} I) > \epsilon$. Therefore, as long as the initial estimate computed in floating point arithmetic $\hat{\eta}^1$ is greater than the exact initial estimate η^1, no singular pencils will be encountered. In particular, the convergence analysis of the previous subsection is valid. Trouble may occur, however, when $\hat{\eta}^1 < \eta^1$ in which case there may not exist θ such that $\sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta} I) > \epsilon$. In general, when $\sigma_{\min}(A - \hat{\eta}^1\,\mathrm{e}^{\mathrm{i}\theta} I) < \epsilon$ for all θ, the circle of radius $\hat{\eta}^1$ lies completely inside the ϵ-pseudospectrum, so the circular search in floating point arithmetic may potentially fail to return any intersection point.

All this discussion suggests raising the initial estimate $\hat{\eta}^1$ by a tolerance. In the next subsection we show that provided structure preserving backward stable eigenvalue solvers are used for the radial searches, $\hat{\eta}^1$ is the imaginary part of the largest imaginary eigenvalue of $K(\theta_\rho, \epsilon + \beta)$, where $\beta = \mathrm{O}(\delta_{\mathrm{mach}}(\|A\| + \epsilon))$. Here δ_{mach} refers to the machine precision and $\mathrm{O}(\delta_{\mathrm{mach}})$ means 'of the order of the machine precision', a standard terminology that is made precise, e.g. in Trefethen & Bau (1997). We have (see Theorem 2.5 and Corollary 2.6)

$$\sigma_{\min}(A - \hat{\eta}^1\,\mathrm{e}^{\mathrm{i}\theta_\rho} I) = \epsilon + \beta.$$

We essentially want to increment $\hat{\eta}^1$ by a value δr such that

$$\sigma_{\min}(A - (\hat{\eta}^1 + \delta r)\,\mathrm{e}^{\mathrm{i}\theta_\rho} I) > \epsilon. \tag{2.16}$$

In a numerical implementation of Algorithm 2.8, the case we need to worry about is when β is negative. Assuming that the multiplicity of $\sigma_{\min}(A - \hat{\eta}^1\,\mathrm{e}^{\mathrm{i}\theta_\rho} I)$ is one (so that Theorem 2.3 implies $\sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta_\rho} I)$ is real-analytic at $r = \hat{\eta}^1$), it follows from the equality

$$(\epsilon + \beta) + \delta r \left.\frac{\partial \sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta_\rho} I)}{\partial r}\right|_{r=\hat{\eta}^1} + \mathrm{O}(\delta r^2) = \sigma_{\min}(A - (\hat{\eta}^1 + \delta r)\,\mathrm{e}^{\mathrm{i}\theta_\rho} I),$$

that for $\delta r = -\beta \Big/ \frac{\partial \sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta_\rho} I)}{\partial r}|_{r=\hat{\eta}^1}$, $\sigma_{\min}(A - (\hat{\eta}^1 + \delta r)\,\mathrm{e}^{\mathrm{i}\theta_\rho} I) = \epsilon + \mathrm{O}(\delta r^2)$ holds. Since according to Theorem 2.3, $\frac{\partial \sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta_\rho} I)}{\partial r}|_{r=\hat{\eta}^1} = \frac{\partial p(r,\theta_\rho)}{\partial r}|_{r=\hat{\eta}^1} = -\mathrm{Re}(\mathrm{e}^{\mathrm{i}\theta_\rho} u^* v)$, where u and v are unit left and right singular vectors corresponding to the minimum singular value of $A - \hat{\eta}^1\,\mathrm{e}^{\mathrm{i}\theta_\rho} I$, we keep incrementing $\hat{\eta}^1$ by $\frac{\beta}{\mathrm{Re}(\mathrm{e}^{\mathrm{i}\theta_\rho} u^* v)}$ until (2.16) is satisfied. Usually it is sufficient to iterate once or twice to obtain a satisfactory δr.

2.6 *Accuracy of Algorithm 2.8*

We analyze the error introduced by a numerical implementation of Algorithm 2.8 that generates increasing estimates in floating point arithmetic and terminates when the circular search fails to return any intersection point. The pseudospectral radius problem (2.5) may be ill conditioned. This is the case when the pseudospectral radius is differentiable and the smallest left and right singular vectors at the global maximizer are close to being orthogonal (see Theorem 2.4). Therefore, we focus on the backward error.

We start with an error analysis for the radial search. From Corollary 2.6 we know that the exact value $\eta_\epsilon(\theta) = r_\epsilon(\theta)$, where $r_\epsilon(\theta)\mathrm{i}$ is the imaginary eigenvalue of $K(\theta,\epsilon)$ with the largest imaginary part. On the other hand, assuming that the eigenvalues are computed by a backward stable algorithm, the counterpart of $r_\epsilon(\theta)$ in floating point arithmetic, say $\hat{r}_\epsilon(\theta)$, is the largest imaginary part of the imaginary eigenvalues of a perturbed matrix

$$\widetilde{K}(\theta,\epsilon) = K(\theta,\epsilon) + E, \tag{2.17}$$

where $\|E\| = \mathrm{O}(\delta_{\mathrm{mach}}\|K(\theta,\epsilon)\|)$ or, since $\|K(\theta,\epsilon)\| \leqslant 2(\|A\| + \epsilon)$, $\|E\| = \mathrm{O}(\delta_{\mathrm{mach}}(\|A\| + \epsilon))$. Additionally, when the algorithm used to solve the Hamiltonian eigenvalue problem is structure preserving, the matrices E and $\widetilde{K}(\theta,\epsilon)$ are Hamiltonian. The analysis for the radial search is valid only when a backward stable, structure preserving Hamiltonian eigenvalue solver (see Chu *et al.*, 2004) is used within Algorithm 2.8.

We first derive an upper bound on the result returned by the radial search in terms of the radius of nearby pseudospectra. The following result inspired by Byers (1988) relates the eigenvalues of $\widetilde{K}(\theta,\epsilon)$ and the $(\epsilon+\beta)$-pseudospectrum of A, where β is some real number with $|\beta|$ at most the norm of the perturbation matrix E.

THEOREM 2.13 (ACCURACY OF THE RADIAL SEARCH) Suppose the Hamiltonian matrix $\widetilde{K}(\theta,\epsilon)$ has the imaginary eigenvalue $\mathrm{i}r$. Then $\mathrm{i}r \in \Lambda(K(\theta,\epsilon+\beta))$ for some real β such that $|\beta| \leqslant \|E\|$.

Proof. Since $\mathrm{i}r \in \Lambda(\widetilde{K}(\theta,\epsilon))$,

$$\det(\widetilde{K}(\theta,\epsilon) - \mathrm{i}rI) = \det(J\widetilde{K}(\theta,\epsilon) - \mathrm{i}rJ) = 0.$$

Notice that $J\widetilde{K}(\theta,\epsilon) - \mathrm{i}rJ$ is Hermitian, meaning that the perturbed Hermitian matrix $J\widetilde{K}(\theta,\epsilon) - \mathrm{i}rJ - JE = JK(\theta,\epsilon) - \mathrm{i}rJ$ has a real eigenvalue β which is at most $\|E\|$ in absolute value (from Weyl's Theorem; see e.g. Horn & Johnson, 1985, Theorem 4.3.1). Now by the definition of $K(\theta,\epsilon)$ (see (2.7))

$$0 = \det(JK(\theta,\epsilon) - \mathrm{i}rJ - \beta I) = \det(K(\theta,\epsilon) + \beta J - \mathrm{i}rI) = \det(K(\theta,\epsilon+\beta) - \mathrm{i}rI).$$

Hence, $\mathrm{i}r$ is an eigenvalue of $K(\theta,\epsilon+\beta)$. □

An immediate consequence of Theorem 2.13 is that $\hat{r}_\epsilon(\theta) \leqslant \eta_{\epsilon+\beta}(\theta)$ for some β with $|\beta| \leqslant \|E\|$; therefore, the result of the radial search in floating point arithmetic, $\hat{r}_\epsilon(\theta)$, satisfies the inequality

$$\hat{r}_\epsilon(\theta) \leqslant \rho_{\epsilon+\|E\|}(A). \tag{2.18}$$

We now turn our attention to the circular search. In order to find the intersection points of the circle of radius r and the ϵ-pseudospectrum boundary, we compute the eigenvalues of the pencil $P(r,\epsilon) - \lambda Q(r,\epsilon)$. In floating point arithmetic, assuming a backward stable algorithm is used, we retrieve the eigenvalues of a nearby pencil $\widetilde{P}(r,\epsilon) - \lambda\widetilde{Q}(r,\epsilon)$. Additionally, for any non-negative real μ, we make use of the notation

$$\widetilde{P}(r,\mu) = P(r,\mu) + E_1 \quad \text{and} \quad \widetilde{Q}(r,\mu) = Q(r,\mu) + E_2,$$

where $E_1 = \widetilde{P}(r,\epsilon) - P(r,\epsilon)$ and $E_2 = \widetilde{Q}(r,\epsilon) - Q(r,\epsilon)$. The fact that eigenvalue solver is backward stable implies $\|E_1\| = \mathrm{O}(\delta_{\mathrm{mach}}(\|A\| + \epsilon + \rho_\epsilon(A)))$ and $\|E_2\| = \mathrm{O}(\delta_{\mathrm{mach}}(\|A\| + \epsilon + \rho_\epsilon(A)))$, since $\|E_1\| = \mathrm{O}(\delta_{\mathrm{mach}}\|P(r,\epsilon)\|)$ and $\|P(r,\epsilon)\| \leqslant (\|A\| + \epsilon + \rho_\epsilon(A))$ and similarly for $\|E_2\|$. The error analysis for the circular search involves the unitary matrix $D(\theta)$ (see (2.10)). The role of $D(\theta)$ in the analysis below is analogous to the role of J in the error analysis for the radial search in the sense that $D(\theta)(P(r,\epsilon) - \mathrm{e}^{\mathrm{i}\theta}Q(r,\epsilon))$ is Hermitian for all θ. In addition to the backward stability requirement on the generalized eigenvalue solver, we also assume it preserves the structure so that for all θ,

$D(\theta)(\widetilde{P}(r,\epsilon)-\mathrm{e}^{\mathrm{i}\theta}\widetilde{Q}(r,\epsilon))$ is Hermitian. Unfortunately, we are not aware of the existence of a backward stable algorithm preserving this structure at the moment, but the assumption that the eigenvalue solver preserves this structure is essential for the analysis.

We are interested in bounding the estimate for the pseudospectral radius from below in terms of a nearby pseudospectral radius when the circular search does not return any intersection point. In this case the pencil $\widetilde{P}(r,\epsilon)-\lambda\widetilde{Q}(r,\epsilon)$ does not have any unit eigenvalue.

THEOREM 2.14 (ACCURACY WHEN THE CIRCULAR SEARCH FAILS) Suppose the pencil $\widetilde{P}(r,\epsilon)-\lambda\widetilde{Q}(r,\epsilon)$ does not have any unit eigenvalue and there exists θ such that $\sigma_{\min}(A-r\,\mathrm{e}^{\mathrm{i}\theta}I)\geqslant\epsilon+\|E_1\|+\|E_2\|$. Then the pencil $P(r,\mu)-\lambda Q(r,\mu)$ does not have any unit eigenvalue for all positive $\mu\leqslant\epsilon-\|E_1\|-\|E_2\|$.

Proof. Let $\chi_j(\theta)$ and $\varphi_j(\theta)$, $j=1,\ldots,2n$, denote the eigenvalues of $N(\theta)=D(\theta)(\widetilde{P}(r,0)-\mathrm{e}^{\mathrm{i}\theta}\widetilde{Q}(r,0))$ and $R(\theta)=D(\theta)(P(r,0)-\mathrm{e}^{\mathrm{i}\theta}Q(r,0))$ as functions of θ in descending order. Notice that $\chi_j(\theta)$ and $\varphi_j(\theta)$ are real-valued continuous functions of θ, since the entries of the matrices $N(\theta)$ and $R(\theta)$ are continuous with respect to θ and both of the matrices are Hermitian for all θ. Note also that

$$|\chi_j(\theta)-\varphi_j(\theta)|\leqslant\|E_1\|+\|E_2\| \tag{2.19}$$

holds for all j and θ. This inequality follows from the fact that $\|N(\theta)-R(\theta)\|=\|E_1-\mathrm{e}^{\mathrm{i}\theta}E_2\|\leqslant\|E_1\|+\|E_2\|$, so the corresponding eigenvalues of the Hermitian matrices cannot differ by more than $\|E_1\|+\|E_2\|$.

Since the pencil $\widetilde{P}(r,\epsilon)-\lambda\widetilde{Q}(r,\epsilon)$ does not have any unit eigenvalue, for all θ

$$\begin{aligned}\det(D(\theta)(\widetilde{P}(r,\epsilon)-\mathrm{e}^{\mathrm{i}\theta}\widetilde{Q}(r,\epsilon)))&=\det(D(\theta)(\widetilde{P}(r,0)-\mathrm{e}^{\mathrm{i}\theta}\widetilde{Q}(r,0)-\epsilon D^*(\theta)))\\&=\det(D(\theta)(\widetilde{P}(r,0)-\mathrm{e}^{\mathrm{i}\theta}\widetilde{Q}(r,0))-\epsilon I)\\&=\det(N(\theta)-\epsilon I)\\&\neq 0.\end{aligned}$$

Hence, the function $\chi_j(\theta)\neq\epsilon$ for all j and θ. But the assumption that $\sigma_{\min}(A-r\,\mathrm{e}^{\mathrm{i}\hat{\theta}}I)\geqslant\epsilon+\|E_1\|+\|E_2\|$ for some $\hat{\theta}$ implies that for all $1\leqslant j\leqslant n$

$$\varphi_j(\hat{\theta})\geqslant\epsilon+\|E_1\|+\|E_2\|, \tag{2.20}$$

since for all θ the eigenvalues of $R(\theta)$ consist of plus and minus the singular values of $A-r\,\mathrm{e}^{\mathrm{i}\theta}I$. When we combine (2.19) and (2.20), we see that for all $1\leqslant j\leqslant n$

$$\chi_j(\hat{\theta})\geqslant\epsilon. \tag{2.21}$$

Now by way of contradiction, suppose there exists a positive $\mu\leqslant\epsilon-\|E_1\|-\|E_2\|$ such that the pencil $P(r,\mu)-\lambda Q(r,\mu)$ has a unit eigenvalue, say $\mathrm{e}^{\mathrm{i}\tilde{\theta}}$. Then

$$\det(D(\tilde{\theta})(P(r,\mu)-\mathrm{e}^{\mathrm{i}\tilde{\theta}}Q(r,\mu)))=\det(D(\tilde{\theta})(P(r,0)-\mathrm{e}^{\mathrm{i}\tilde{\theta}}Q(r,0))-\mu I)=0$$

meaning for some $j\leqslant n$, $\varphi_j(\tilde{\theta})=\mu$, since μ is positive. It follows from (2.19) that $\chi_j(\tilde{\theta})\leqslant\mu+\|E_1\|+\|E_2\|\leqslant\epsilon$ is satisfied. But for the same j, (2.21) holds as well. We conclude from the intermediate value theorem that there exists θ' satisfying $\chi_j(\theta')=\epsilon$. This contradicts the fact that $\chi_j(\theta)\neq\epsilon$ for all θ and j. □

In exact arithmetic the circular search fails when the circle of radius r lies either completely inside or completely outside the pseudospectrum. In the theorem above we need the condition that there exists a point $r\,\mathrm{e}^{\mathrm{i}\theta}$ on the circle of radius r such that $\sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta} I) \geqslant \epsilon + \|E_1\| + \|E_2\|$ in order to distinguish these two cases. When $\sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta} I) \geqslant \epsilon$ and the derivative of $\sigma_{\min}(A - r'\,\mathrm{e}^{\mathrm{i}\theta'} I)$ with respect to θ' at (r, θ) is not very small, such a point exists on the circle of radius r in a small neighborhood of θ. In the previous subsection we discussed how to generate estimates r such that $\sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta_\rho} I) \geqslant \epsilon$.

Focusing on the implications of Theorem 2.14, whenever the circular search in floating point arithmetic fails for some $r > \rho(A)$ and there exists a point (r, θ) with $\sigma_{\min}(A - r\,\mathrm{e}^{\mathrm{i}\theta} I) \geqslant \epsilon + \|E_1\| + \|E_2\|$, then for all $\tau \geqslant \|E_1\| + \|E_2\|$, the $(\epsilon - \tau)$-pseudospectrum lies inside the circle of radius r. We infer the lower bound

$$\rho_{\epsilon - \|E_1\| - \|E_2\|}(A) \leqslant r. \tag{2.22}$$

Now we are ready to find the backward error of the algorithm. First, since the estimates are increasing in floating point arithmetic, the algorithm is guaranteed to terminate. At the termination the estimate value must satisfy the upper bound (2.18), because it is generated by a radial search at the previous iteration. Moreover, at the last iteration, the circular search fails meaning that the lower bound (2.22) on the final estimate holds as well. Combining these bounds and from the continuity of $\rho_\epsilon(A)$ with respect to ϵ (see Theorem 2.1), we see that the estimate $\hat{\rho}_\epsilon(A)$ at the termination satisfies

$$\hat{\rho}_\epsilon(A) = \rho_{\epsilon+\beta}(A),$$

where $\beta = \mathrm{O}(\delta_{\mathrm{mach}}(\|A\| + \epsilon + \rho_\epsilon(A)))$, i.e. the final estimate is the solution of a nearby pseudospectral radius problem for the same matrix.

Since our analysis above depends on the usage of the proper eigenvalue solvers, let us comment on the eigenvalue solvers that should ideally be used in a reliable implementation of Algorithm 2.8. For the radial searches inside the current implementation we use the backward stable Hamiltonian eigenvalue solver in Benner *et al.* (1998). Also in Benner *et al.* (1998) a generalized Hamiltonian eigenvalue solver is described. For the circular searches, the symplectic pencil $P(r, \epsilon)^* - \lambda Q(r, \epsilon)^*$ can be converted into a Hamiltonian pencil via a Cayley transformation and the algorithm in Benner *et al.* (1998) can be applied. To be precise the algorithms in Benner *et al.* (1998) preserve the symmetry of the eigenvalues rather than the structure of the matrix. Using eigenvalue solvers preserving the eigenvalue structure is sufficient to avoid the tolerances needed to decide whether an eigenvalue is imaginary or of unit modulus. However, the analysis above does not necessarily apply in this case. Recently, Chu *et al.* (2004) presented backward stable structure preserving algorithms that are the most suitable choices inside Algorithm 2.8 at the moment, but implementations are not yet available.

We note that an analogous backward error analysis applies to the criss-cross method for the pseudospectral abscissa (Burke *et al.*, 2003b) and is available on the web (Mengi & Overton, 2004a). Additionally, on the web site (Mengi & Overton, 2004b) we provide an extreme numerical example for which the radial search introduces a large error that can be explained by the analysis in this section. Remarkably, the errors introduced by the radial searches do not affect the accuracy of the overall algorithm.

3. Computation of the numerical radius

An analogous measure to the numerical radius for continuous-time dynamical systems is the numerical abscissa, the real component of the rightmost point in the field of values,

$$\alpha_F(A) = \max\{\mathrm{Re}\ z \colon z \in F(A)\}.$$

Though intuitively computation of the numerical abscissa and the numerical radius seem equally difficult, the former can be reduced to an eigenvalue problem (Horn & Johnson, 1991)

$$\alpha_F(A) = \lambda_{\max}(H(A)), \tag{3.1}$$

where $H(A) = \frac{1}{2}(A + A^*)$. By contrast, the computation of $r(A)$ is a challenging task.

Multiplying A by $\mathrm{e}^{\mathrm{i}\theta}$ rotates the field of values of A by θ. Consequently, the numerical radius of A can be viewed as the global maximum of an optimization problem with a single real variable

$$r(A) = \max_{\theta\in[0,2\pi)} \alpha_F(A\,\mathrm{e}^{\mathrm{i}\theta}). \tag{3.2}$$

Combining (3.1) and (3.2) yields the following characterization of the numerical radius

$$r(A) = \max_{\theta\in[0,2\pi)} \lambda_{\max}(H(A\,\mathrm{e}^{\mathrm{i}\theta})).$$

Given the matrix A, let us define $f\colon [0, 2\pi) \to \mathbb{R}$ by

$$f(\theta) = \lambda_{\max}(H(A\,\mathrm{e}^{\mathrm{i}\theta})). \tag{3.3}$$

Observe that for each $\theta \in [0, 2\pi)$, $f(\theta) \in [-\|A\|, \|A\|]$. Our aim is to find the global maximum of f.

In our algorithm, we need to determine θ values satisfying $f(\theta) = \alpha$, where $\alpha > 0$ is a numerical radius estimate. Consider the pencil $R(\alpha) - \lambda S$ with

$$R(\alpha) = \begin{bmatrix} 2\alpha I & -A^* \\ I & 0 \end{bmatrix}, \quad S = \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}.$$

In He & Watson (1997), it is proved that given a real number $\alpha \geqslant \min_\theta f(\theta)$, the pencil $R(\alpha) - \lambda S$ has an eigenvalue on the unit circle or is singular if and only if the inequality $\alpha \leqslant r(A)$ holds. Using the theorem in He & Watson (1997), we can decide whether there is a θ satisfying $f(\theta) = \alpha$; however, this theorem does not tell us what the θ values are. For this purpose we state a slightly modified version.

THEOREM 3.1 The pencil $R(\alpha) - \lambda S$ has the eigenvalue $\mathrm{e}^{\mathrm{i}\theta}$ or is singular if and only if the Hermitian matrix $H(A\,\mathrm{e}^{\mathrm{i}\theta})$ has α as one of its eigenvalues.

Proof. The equality $\det(R(\alpha) - \mathrm{e}^{\mathrm{i}\theta} S) = 0$ is satisfied if and only if the matrix

$$\begin{bmatrix} 2\alpha I - \mathrm{e}^{\mathrm{i}\theta} A & -A^* \\ I & -\mathrm{e}^{\mathrm{i}\theta} I \end{bmatrix}$$

is singular. Multiplying the bottom block row of this matrix by $\mathrm{e}^{-\mathrm{i}\theta}$, we see that this $2n \times 2n$ matrix is singular if and only if the $n \times n$ matrix $\mathrm{e}^{\mathrm{i}\theta} A + \mathrm{e}^{-\mathrm{i}\theta} A^* - 2\alpha I$ is singular. Therefore, the Hermitian matrix $H(A\,\mathrm{e}^{\mathrm{i}\theta})$ has α as one of its eigenvalues if and only if the matrix $R(\alpha) - \mathrm{e}^{\mathrm{i}\theta} S$ is rank deficient. □

From Theorem 3.1 it follows that as long as the pencil $R(\alpha) - \lambda S$ is regular for a given α we can solve the generalized eigenvalue problem $R(\alpha) - \lambda S$ and extract the angles of the eigenvalues on the unit circle to obtain a superset of θ values satisfying $f(\theta) = \alpha$. To determine the exact set, for each angle θ' that is extracted, the eigenvalues of $H(A\,\mathrm{e}^{\mathrm{i}\theta'})$ need to be computed. Only those angles for which $H(A\,\mathrm{e}^{\mathrm{i}\theta'})$ has α as the largest eigenvalue should be kept.

Now that we know how to compute the intersection points of a horizontal line with the graph of f efficiently, we suggest an iterative algorithm. At the jth iteration the algorithm generates an estimate of the numerical radius, r^j, and a set of open intervals, $I_1^j, I_2^j, \ldots, I_{m^j}^j$, where, as earlier, $I_{m_j}^j$ may wrap around the circle. The function f is greater than r^j in each interval I_k^j, $1 \leqslant k \leqslant m^j$ (i.e. for all $\theta \in I_k^j$, $f(\theta) > r^j$) and exactly r^j at the end points of the intervals. At the jth iteration the new estimate r^j is set to the maximum value the function f attains at the midpoints of the open intervals produced at the previous iteration. Then the open intervals at the jth iteration are obtained using Theorem 3.1 followed by the maximum eigenvalue checks.

ALGORITHM 3.2

1. Set $j = 0$ and $\phi^0 = [0]$.
2. Update the numerical radius estimate: Compute r^{j+1} using the formula
$$r^{j+1} = \max\{f(\theta) : \theta \in \phi^j\}. \tag{3.4}$$
3. Update the set of the midpoints: Find θ values for which $f(\theta) = r^{j+1}$ holds. From these infer the open intervals $I_1^{j+1}, I_2^{j+1}, \ldots, I_{m^{j+1}}^{j+1}$ in which $f(\theta) > r^{j+1}$. Calculate the new set of midpoints
$$\phi^{j+1} = \{\phi_1^{j+1}, \phi_2^{j+1}, \ldots, \phi_{m^{j+1}}^{j+1}\},$$
where ϕ_k^{j+1} is the midpoint of the open interval I_k^{j+1}
$$\phi_k^{j+1} = \begin{cases} \frac{\gamma_k^{j+1} + \zeta_k^{j+1}}{2}, & \text{if } \gamma_k^{j+1} < \zeta_k^{j+1}, \\ \frac{\gamma_k^{j+1} + \zeta_k^{j+1} + 2\pi}{2} \bmod 2\pi, & \text{otherwise.} \end{cases}$$
4. Increment j by one, go to step 2.

A robust way to determine $I_1^j, I_2^j, \ldots, I_{m^j}^j$ from the set of intersection points is to sort the intersection points and to compute f at the midpoint of each adjacent pair of points. The pencil $R(\alpha)^* - \lambda S^*$ is symplectic, so just as with circular searches in the previous section, this problem can be reduced to a Hamiltonian eigenvalue problem and the eigenvalue solver described in Benner *et al.* (1998) can be applied. It is easy to avoid singular pencils, since for all α greater than the initial estimate $r^1 = f(0)$, the pencil $R(\alpha) - \lambda S$ is guaranteed to be regular. Note also that we can compute f accurately because of the fact that the eigenvalues of symmetric matrices are well conditioned.

Algorithm 3.2 is an extension of the Boyd–Balakrishnan algorithm (Boyd & Balakrishnan, 1990) to the numerical radius. Thus, a similar convergence proof applies (based on the fact that the length of the greatest open interval is at least halved at each iteration). We believe that a proof along the line of the argument in Boyd & Balakrishnan (1990) is applicable to show that the algorithm converges quadratically to the value $r(A)$ and that the accuracy analysis in the previous section for the pseudospectral radius can be extended to Algorithm 3.2.

4. Software

The software package *EigTool* (Wright, 2002) is a MATLAB graphical user interface used to compute the pseudospectra of a matrix. The Numbers menu includes options to compute the ϵ-pseudospectral

abscissa using the algorithm of Burke *et al.* (2003b) and the ϵ-pseudospectral radius using the algorithm presented in this paper. In addition, freely downloadable implementations of these algorithms and others, including the algorithm of Section 3 to compute the numerical radius, may be found at Mengi & Overton (2005).

Acknowledgements

Many thanks to A. S. Lewis for helpful suggestions and conversations. We also would like to thank M. Embree, L. N. Trefethen and an anonymous referee for carefully reading the manuscript and for their constructive comments which greatly improved the quality of this paper. This research was supported in part by National Science Foundation Grant CCR-0098145.

REFERENCES

AXELSSON, O., LU, H. & POLMAN, B. (1994) On the numerical radius of matrices and its application to iterative solution methods. *Linear Multilinear Algebra*, **37**, 225–238.

BENNER, P., MEHRMANN, V. & XU, H. (1998) A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils. *Numer. Math.*, **78**, 329–358.

BONNANS, J. F. & SHAPIRO, A. (2000) *Perturbation Analysis of Optimization Problems*. New York: Springer.

BOYD, S. & BALAKRISHNAN, V. (1990) A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞-norm. *Syst. Control Lett.*, **15**, 1–7.

BURKE, J. V., LEWIS, A. S. & OVERTON, M. L. (2003a) Optimization and pseudospectra, with applications to robust stability. *SIAM J. Matrix Anal.*, **25**, 80–104.

BURKE, J. V., LEWIS, A. S. & OVERTON, M. L. (2003b) Robust stability and a criss-cross algorithm for pseudospectra. *IMA J. Numer. Anal.*, **23**, 359–375.

BYERS, R. (1988) A bisection method for measuring the distance of a stable matrix to the unstable matrices. *SIAM J. Sci. Stat. Comput.*, **9**, 875–881.

CHU, D., LIU, X. & MEHRMANN, V. (2004) A numerically strongly stable method for computing the Hamiltonian Schur form. *Technical Report 24/2004*. Berlin: Institut für Mathematik, Technische Universität Berlin.

DEMMEL, J. W. (1987) A counterexample for two conjectures about stability. *IEEE Trans. Auto. Control*, **32**, 340–342.

EIERMANN, M. (1993) Field of values & iterative methods. *Linear Algebra Appl.*, **180**, 167–197.

EMBREE, M. & TREFETHEN, L. N. (2005) Pseudospectra gateway. http://web.comlab.ox.ac.uk/projects/pseudospectra/.

HE, C. & WATSON, G. A. (1997) An algorithm for computing the numerical radius. *IMA J. Numer. Anal.*, **17**, 329–342.

HORN, R. A. & JOHNSON, C. R. (1985) *Matrix Analysis*. Cambridge: Cambridge University Press.

HORN, R. A. & JOHNSON, C. R. (1991) *Topics in Matrix Analysis*. Cambridge: Cambridge University Press.

KREISS, H.-O. (1962) Über die Stabilitätsdefinition für Differenzengleichungen die partielle Differenzialgleichungen approximieren. *BIT*, **2**, 153–181.

LEWIS, A. S. (2002) Robust regularization. *Technical Report*, Simon Fraser University. (submitted to Mathematical Programming).

MENGI, E. & OVERTON, M. L. (2004a) Backward error analysis for the criss-cross method for the pseudospectral abscissa. http://www.cs.nyu.edu/mengi/robust_stability/berrorpseudoabs.ps.

MENGI, E. & OVERTON, M. L. (2004b) An extreme example about the accuracy of the radial search. http://www.cs.nyu.edu/mengi/robust_stability/extexample.html.

MENGI, E. & OVERTON, M. L. (2005) Software for robust stability measures. http://www.cs.nyu.edu/mengi/robuststability.html.

PEARCY, C. (1966) An elementary proof of the power inequality for the numerical radius. *Mich. Math. J.*, **13**, 284–291.

TREFETHEN, L. N. & BAU III, D. (1997) *Numerical Linear Algebra.* Philadelphia: SIAM.

TREFETHEN, L. N. & EMBREE, M. (2005) *Spectra & Pseudospectra: The Behavior of Nonnormal Matrices*. Princeton: Princeton University Press.

WEGERT, E. & TREFETHEN, L. N. (1994) From the Buffon needle problem to the Kreiss matrix. *Am. Math. Mon.*, **101**, 132–139.

WRIGHT, T. G. (2002) EigTool: a graphical tool for nonsymmetric eigenproblems. Oxford University Computing Labora- tory. http://www.comlab.ox.ac.uk/pseudospectra/eigtool/.