SIAM J. MATRIX ANAL. APPL.
Vol. 25, No. 1, pp. 80–104

OPTIMIZATION AND PSEUDOSPECTRA, WITH APPLICATIONS TO ROBUST STABILITY*

J. V. BURKE†, A. S. LEWIS‡, AND M. L. OVERTON§

Abstract. The ϵ-pseudospectrum of a matrix A is the subset of the complex plane consisting of all eigenvalues of all complex matrices within a distance ϵ of A. We are interested in *two* aspects of "optimization and pseudospectra." The first concerns *maximizing* the function "real part" over an ϵ-pseudospectrum of a fixed matrix: this defines a function known as the ϵ-pseudospectral abscissa of a matrix. We present a bisection algorithm to compute this function. Our second interest is in *minimizing* the ϵ-pseudospectral abscissa over a set of feasible matrices. A prerequisite for local optimization of this function is an understanding of its variational properties, the study of which is the main focus of the paper. We show that, in a neighborhood of any nonderogatory matrix, the ϵ-pseudospectral abscissa is a nonsmooth but locally Lipschitz and subdifferentially regular function for sufficiently small ϵ; in fact, it can be expressed locally as the maximum of a finite number of smooth functions. Along the way we obtain an eigenvalue perturbation result: near a nonderogatory matrix, the eigenvalues satisfy a Hölder continuity property on matrix space—a property that is well known when only a single perturbation parameter is considered. The pseudospectral abscissa is a powerful modeling tool: not only is it a robust measure of stability, but it also reveals the transient (as opposed to asymptotic) behavior of associated dynamical systems.

Key words. pseudospectrum, eigenvalue optimization, spectral abscissa, nonsmooth analysis, subdifferential regularity, robust optimization, robust control, stability radius, distance to instability, $\mathbf{H}_\infty$ norm

AMS subject classifications. Primary, 15A18, 65K05; Secondary, 90C30, 93D09

PII. S0895479802402818

1. Introduction.

The ϵ-pseudospectrum of a matrix A, denoted $\Lambda_\epsilon(A)$, is the subset of the complex plane consisting of all eigenvalues of all complex matrices within a distance ϵ of A (see [20, 39, 40]). We are interested in *two* aspects of "optimization and pseudospectra." The first concerns *maximizing* a simple real-valued function over a fixed pseudospectrum $\Lambda_\epsilon(A)$. We focus specifically on the case where this function is simply "real part." Then the optimal value defines the ϵ-pseudospectral abscissa of A, denoted $\alpha_\epsilon(A)$. Just as the spectral abscissa of a matrix provides a measure of its stability, that is, the asymptotic decay of associated dynamical systems, so the ϵ-pseudospectral abscissa provides a measure of *robust stability*, where by robust we mean with respect to complex perturbations in the matrix. One of the contributions of this paper is a bisection algorithm that computes $\alpha_\epsilon(A)$ for any A; this algorithm also identifies all maximizing points in the pseudospectrum.

In many applications, matrices are not fixed but dependent on parameters that may be adjusted. Our second interest in optimization concerns *minimizing* the ϵ-pseudospectral abscissa α_ϵ over a feasible set of matrices. A prerequisite for local

*Received by the editors February 15, 2002; accepted for publication (in revised form) by L. Vandenberghe January 3, 2003; published electronically May 15, 2003.

http://www.siam.org/journals/simax/25-1/40281.html

†Department of Mathematics, University of Washington, Seattle, WA 98195 (burke@math.washington.edu). The research of this author was supported in part by National Science Foundation grant DMS-9971852.

‡Department of Mathematics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada (aslewis@sfu.ca, www.cecm.sfu.ca/~aslewis). The research of this author was supported in part by NSERC.

§Courant Institute of Mathematical Sciences, New York University, New York, NY 10012 (overton@cs.nyu.edu). The research of this author was supported in part by National Science Foundation grant CCR-0098145.

minimization of α_ϵ is an understanding of its variational properties as a function of the matrix A. This provides the focus for most of the paper. Our main result shows that, in a neighborhood of any nonderogatory matrix, the ϵ-pseudospectral abscissa is a nonsmooth but locally Lipschitz and subdifferentially regular function for sufficiently small ϵ; in fact, it can be expressed locally as the maximum of a finite number of smooth functions. Such a property is desirable from the point of view of numerical methods for local optimization, but we defer a computational study to future work.

The paper is organized as follows. After setting up some notation in section 2, we begin in section 3 by discussing related ideas in the robust control literature. We review the connections between the pseudospectral abscissa and the "distance to instability" [28, 16], or "complex stability radius" [21], and the $\mathbf{H}_\infty$ norm of a transfer function [8]. The outcome of minimization of α_ϵ over a set of matrices obviously depends on the crucial issue of the choice of ϵ. We show that as ϵ is increased from zero to an arbitrarily large quantity, the corresponding optimization problem evolves from minimization of the spectral abscissa (enhancing the *asymptotic* decay rate of the associated dynamical system) to the minimization of the largest eigenvalue of the symmetric part of the matrix (minimizing the *initial* growth rate of the associated system). Regarding the first of these extremes (optimization of the spectral abscissa), variational analysis of this non-Lipschitz function is well understood [15, 12], global optimization is known to be hard [5], and some progress has been made in local optimization methods [14]. Regarding the second extreme, optimization of this convex function over a polyhedral feasible set is a semidefinite programming problem, and the global minimum can be found by standard methods [4, 38]. Intermediate choices of ϵ control *transient* peaking in the dynamical system associated with the matrix, and one particular choice corresponds exactly to the complex stability radius (or $\mathbf{H}_\infty$ norm) optimization problem. Thus the pseudospectral approach gives a whole range of stabilizing optimization problems, each with a quantifiable interpretation in terms of the allowable perturbations. Furthermore, unlike maximization of the complex stability radius, which simply optimizes the "robustness" of the stability, minimizing the pseudospectral abscissa preserves some explicit emphasis on optimizing the asymptotic decay rate of the system.

In section 4, we analyze the topology of the pseudospectrum, observing that points on the boundary are accessible from the interior by analytic paths, and discussing conditions under which the boundary is differentiable at points that maximize the real part. This sets the stage for the description of a simple bisection algorithm to compute $\alpha_\epsilon(A)$, the pseudospectral abscissa of a fixed matrix, in section 5. In section 6, we show that the bisection algorithm locates *all* maximizers of the real part over the pseudospectrum. The bisection algorithm is very much analogous to Byers' algorithm for measuring the distance to instability [16], which has spawned more sophisticated variants for the calculation of stability radii (real as well as complex) and $\mathbf{H}_\infty$ norms, both globally and quadratically (or higher order) convergent; see [8, 7, 11, 18, 36]. Along similar lines, we have also developed a quadratically convergent variant algorithm for computing α_ϵ, described and analyzed in a companion paper [13].

In section 7, we continue our study of analytical properties of the pseudospectrum. As is well known, the pseudospectrum of a matrix A is defined by an inequality on $\sigma_{\min}(A - zI)$, the least singular value of $A - zI$. In Theorem 7.4 (growth near an eigenvalue) we give an interesting estimate relating $\sigma_{\min}(A - zI)$ to $|z - \lambda_0|^m$, where

λ_0 is a nonderogatory eigenvalue (one whose geometric multiplicity is one), and m is its algebraic multiplicity. The coefficient relating these quantities is a ratio of two products, of eigenvalue separations and of singular values, respectively. One corollary of this result is that an ϵ-pseudospectrum component around a nonderogatory eigenvalue is strictly convex for sufficiently small ϵ, an intuitively appealing but apparently nontrivial fact. Another corollary is that the nonderogatory eigenvalue λ_0 satisfies a Hölder continuity property in a neighborhood of A in matrix space, with Hölder exponent equal to $1/m$. While this result might not seem surprising, in light of well-known classical spectral perturbation theory [27, 23, 3, 30], we have not seen it in the literature. The classical analysis focuses almost exclusively on single perturbation parameters.

The analytical results of section 7 allow us to achieve our primary goal in section 8: a detailed variational analysis of the pseudospectral abscissa α_ϵ. The main result has already been mentioned above. Finally, in section 9, we examine the boundary properties of the pseudospectrum at points where the boundary is not smooth, using techniques from modern variational analysis [17, 34]. We show that, under a nondegeneracy condition, the complement of the pseudospectrum is Clarke regular at such a point, and give a formula for the normal cone.

2. Notation. We consider a matrix A in the space of $n \times n$ complex matrices $\mathbf{M}^n$. We denote the spectrum of A by $\Lambda = \Lambda(A)$, and we denote by $\alpha = \alpha(A)$ the *spectral abscissa* of A, which is the largest of the real parts of the eigenvalues.

For a real $\epsilon > 0$, the *ϵ-pseudospectrum* of A is the set

$$\Lambda_\epsilon = \{z \in \mathbf{C} : z \in \Lambda(X) \text{ where } \|X - A\| \le \epsilon\}.$$

(Throughout, $\|\cdot\|$ denotes the operator 2-norm on $\mathbf{M}^n$.) For the most part, ϵ is fixed, so where it is understood we drop it from the terminology. Any element of the pseudospectrum is called a *pseudoeigenvalue*. Unless otherwise stated, we shall always assume $\epsilon > 0$, but it is occasionally helpful to extend our notation to allow $\epsilon = 0$, so $\Lambda_0 = \Lambda$. Analogously, the *strict pseudospectrum* is the set

$$\Lambda'_\epsilon = \{z \in \mathbf{C} : z \in \Lambda(X) \text{ where } \|X - A\| < \epsilon\}.$$

The *pseudospectral abscissa* α_ϵ is the maximum value of the real part over the pseudospectrum:

$$\alpha_\epsilon = \sup\{\operatorname{Re} z : z \in \Lambda_\epsilon\}. \tag{2.1}$$

We call this optimization problem *the pseudospectral abscissa problem*. Note $\alpha_0 = \alpha$.

The function $\sigma_{\min} : \mathbf{M}^n \to \mathbf{R}$ denotes the smallest singular value. We define a function $g : \mathbf{C} \to \mathbf{R}$ by

$$g(z) = \sigma_{\min}(A - zI) = \|(A - zI)^{-1}\|^{-1},$$

where we interpret the right-hand side as zero when $z \in \Lambda(A)$. Thus g is the reciprocal of the norm of the resolvent. Using this notation, a useful characterization of the pseudospectrum is

$$\Lambda_\epsilon = \{z \in \mathbf{C} : g(z) \le \epsilon\},$$

and analogously

$$\Lambda'_\epsilon = \{z \in \mathbf{C} : g(z) < \epsilon\}$$

(see [39]). Clearly as ϵ increases, both families of sets are monotonic increasing.

We will sometimes want to allow the matrix A (and the parameter ϵ) to vary. We therefore define the pseudospectral abscissa *function* $\alpha_\epsilon : \mathbf{M}^n \to \mathbf{R}$ by

$$\alpha_\epsilon(Z) = \sup\{\operatorname{Re} z : \sigma_{\min}(Z - zI) \le \epsilon\}.$$

3. Related ideas. The pseudospectral abscissa is related to several other functions important for stability analysis. In this section we briefly sketch the connections with two such functions, in particular, the "distance to instability" and the $\mathbf{H}_\infty$ norm.

A matrix A is *stable* if all its eigenvalues have strictly negative real parts; in other words, the spectral abscissa of A satisfies $\alpha(A) < 0$. From any given matrix A, the distance to the set of matrices which are not stable [28, 19] (also known as the *complex stability radius* [21]) is

$$\beta(A) = \min\{\|X - A\| : X \in \mathbf{M}^n,\ \alpha(X) \ge 0\}.$$

Since the set of matrices which are not stable is closed, this minimum is attained. Notice in particular that $\beta(A) = 0$ if and only if A is not stable. It is now easy to check the relationship

$$\beta(A) \le \epsilon \quad \Leftrightarrow \quad \alpha_\epsilon(A) \ge 0, \tag{3.1}$$

and more generally, for any real x,

$$\alpha_\epsilon(A) \ge x \quad \Leftrightarrow \quad \alpha_\epsilon(A - xI) \ge 0 \quad \Leftrightarrow \quad \beta(A - xI) \le \epsilon.$$

Notice that we can write the pseudospectral abscissa in the form

$$\alpha_\epsilon(A) = \max\{\alpha(X) : \|X - A\| \le \epsilon\},$$

a special case of "robust regularization" [26] and "minimum stability degree" [2]. Since the spectral abscissa α is continuous, standard arguments [26] show that the function

$$(\epsilon, A) \in \mathbf{R}_+ \times \mathbf{M}^n \mapsto \alpha_\epsilon(A) \tag{3.2}$$

is continuous.

In this paper we consider almost exclusively a fixed choice of the parameter ϵ, but for the moment let us consider the effect of varying ϵ on the solution of a pseudospectral abscissa minimization problem. For any fixed set of feasible matrices $F \subset \mathbf{M}^n$, the continuity of the map (3.2) guarantees various useful continuity properties of the optimal value and solutions of the optimization problem $\inf_F \alpha_\epsilon$ (see [34, Chap. 7]). In particular, if F is nonempty and compact, then

$$\lim_{\epsilon \to \bar\epsilon} \inf_F \alpha_\epsilon = \inf_F \alpha_{\bar\epsilon},$$

and any cluster point of a sequence of matrices A_r minimizing α_{ϵ_r} over F, where $\epsilon_r \to \bar\epsilon$, must minimize $\alpha_{\bar\epsilon}$ over F.

Notice that any stable matrix A satisfies

$$\alpha_{\beta(A)}(A) = 0.$$

To see this, note that the implication (3.1) shows $\alpha_{\beta(A)} \ge 0$, while if $\alpha_{\beta(A)} > 0$, then by the continuity of α_ϵ with respect to ϵ, there would exist $\epsilon \in (0, \beta(A))$ such that $\alpha_\epsilon(A) \ge 0$, whence we get the contradiction $\beta(A) \le \epsilon < \beta(A)$.

We return to our pseudospectral abscissa minimization problem $\inf_F \alpha_\epsilon$. The following easy result shows that, under reasonable conditions, for a particular choice of ϵ, this problem is equivalent to maximizing the distance to instability over the same set of feasible matrices.

PROPOSITION 3.1 (maximizing the distance to instability). *If the optimal value* $\bar\beta = \max_F \beta$ *is attained by some stable matrix, then* $\min_F \alpha_{\bar\beta} = 0$ *and*

$$\operatorname{argmin}\{\alpha_{\bar\beta}(X) : X \in F\} = \operatorname{argmax}\{\beta(X) : X \in F\}.$$

Proof. Any matrix $A \in F$ satisfies $\beta(A) \le \bar\beta$. If A is stable, then $\alpha_{\bar\beta}(A) \ge \alpha_{\beta(A)}(A) = 0$, while on the other hand, if A is not stable, then $\alpha_{\bar\beta}(A) \ge \alpha_0(A) \ge 0$. Hence $\inf_F \alpha_{\bar\beta} \ge 0$.

By assumption, $\bar\beta$ is finite and strictly positive, so clearly every matrix in the (nonempty) set of optimal solutions $\operatorname{argmax}_F \beta$ is stable. Any such matrix A satisfies $\alpha_{\bar\beta}(A) = \alpha_{\beta(A)}(A) = 0$, and hence $A \in \operatorname{argmin}_F \alpha_{\bar\beta}$. We deduce $\operatorname{argmax}_F \beta \subset \operatorname{argmin}_F \alpha_{\bar\beta}$ and $\min_F \alpha_{\bar\beta} = 0$.

Consider, conversely, a matrix $A \in F$ such that $A \notin \operatorname{argmax}_F \beta$. Suppose first that A is stable. Since $\beta(A) < \bar\beta$, we know $\alpha_{\bar\beta}(A) > \alpha_{\beta(A)}(A) = 0$, because as we shall see in the next section, $\alpha_\epsilon(A)$ is strictly increasing in ϵ. On the other hand, if A is not stable, then the same reasoning shows $\alpha_{\bar\beta}(A) > \alpha_0(A) \ge 0$. In either case, we have shown $A \notin \operatorname{argmin}_F \alpha_{\bar\beta}$, so $\operatorname{argmin}_F \alpha_{\bar\beta} \subset \operatorname{argmax}_F \beta$ as required. ∎

We thus see that, under reasonable conditions, as ϵ increases from zero, the set of optimal solutions $\operatorname{argmax}_F \alpha_\epsilon$ evolves from the set of minimizers of the spectral abscissa through the set of maximizers of the stability radius. This raises the question of what happens for *large* ϵ. The following result shows that the limiting version of $\inf_F \alpha_\epsilon$ as $\epsilon \to +\infty$ is the optimization problem

$$\inf_{X \in F} \lambda_{\max}\Big(\frac{X + X^*}{2}\Big),$$

where $\lambda_{\max}$ denotes the largest eigenvalue of a Hermitian matrix.

THEOREM 3.2 (large ϵ). *For any matrix* $A \in \mathbf{M}^n$,

$$\big[\alpha_\epsilon(X) - \epsilon\big] \to \lambda_{\max}\Big(\frac{A + A^*}{2}\Big) \quad \text{as} \quad \epsilon \to +\infty \text{ and } X \to A.$$

Proof. If we denote the right-hand side by λ, then there is a unit vector $u \in \mathbf{C}^n$ satisfying $u^*(A + A^*)u = 2\lambda$. Consider any sequence $\epsilon_r \to +\infty$ and $X_r \to A$. Since $\|uu^*\| = 1$, we know

$$\alpha_{\epsilon_r}(X_r) - \epsilon_r \ge \alpha(X_r + \epsilon_r uu^*) - \epsilon_r = \epsilon_r\Big(\alpha\Big(uu^* + \frac{1}{\epsilon_r}X_r\Big) - 1\Big).$$

Now standard perturbation theory [23] shows α is analytic around the matrix uu^* with gradient $\nabla\alpha(uu^*) = uu^*$, so as $r \to \infty$, the right-hand side in the above relationship converges to

$$\operatorname{Re}(\operatorname{tr}(uu^*A)) = \operatorname{Re} u^*Au = \lambda.$$

We have thus shown

$$\liminf_r (\alpha_{\epsilon_r}(X_r) - \epsilon_r) \ge \lambda.$$

Now suppose

$$\limsup_r (\alpha_{\epsilon_r}(X_r) - \epsilon_r) > \lambda.$$

We will derive a contradiction. Without loss of generality, there exists a real $\delta > 0$ such that

$$\alpha_{\epsilon_r}(X_r) - \epsilon_r > \lambda + \delta \quad \text{for all } r.$$

For each r we can choose a matrix D_r satisfying $\|D_r\| \leq 1$ and

$$\alpha_{\epsilon_r}(X_r) = \alpha(X_r + \epsilon_r D_r),$$

and a unit vector $w_r \in \mathbf{C}^n$ satisfying

$$\alpha(X_r + \epsilon_r D_r) = \mathrm{Re}\,(w_r^*(X_r + \epsilon_r D_r)w_r).$$

Hence

$$\begin{aligned}\lambda + \delta &< \mathrm{Re}\,(w_r^* X_r w_r) + \epsilon_r(\mathrm{Re}\,(w_r^* D_r w_r) - 1)\\ &\leq \mathrm{Re}\,(w_r^* X_r w_r) = w_r^*\Big(\frac{X_r + X_r^*}{2}\Big)w_r\\ &\leq \lambda_{\max}\Big(\frac{X_r + X_r^*}{2}\Big).\end{aligned}$$

But as $r \to \infty$, the right-hand side above converges to λ, which is the desired contradiction. ◻

We see from this result that, for example, if the set F is a polyhedron, then the limiting version of the optimization problem $\inf_F \alpha_\epsilon$ as $\epsilon \to \infty$ is a computationally straightforward, convex minimization problem, whereas when $\epsilon = 0$ the problem may be hard [5].

The idea of the $\mathbf{H}_\infty$ norm of a transfer matrix is also closely related to the complex stability radius. Consider the linear time-invariant dynamical system

$$\dot{p} = Ap + u,$$

where p denotes the state vector (in this simple case coinciding with the output) and u denotes the input vector. The "transfer matrix" of this system is the function $H(s) = (sI - A)^{-1}$ (where s is a complex variable). Assuming the matrix A is stable, the corresponding $\mathbf{H}_\infty$ norm is defined by

$$\|H\|_\infty = \sup_{\omega \in \mathbf{R}} \sigma_{\max}(H(i\omega)),$$

where $\sigma_{\max}$ denotes the largest singular value. Clearly

$$\|H\|_\infty = \sup_{\omega \in \mathbf{R}} \frac{1}{\sigma_{\min}(A - i\omega I)},$$

so $\|H\|_\infty < \epsilon^{-1}$ if and only if we have

$$\sigma_{\min}(A - i\omega I) > \epsilon \quad \text{for all } \omega \in \mathbf{R}.$$

As a consequence of Theorem 5.4 below, for example, this is equivalent to $\alpha_\epsilon(A) < 0$. In summary, for a stable matrix A, we have

$$\alpha_\epsilon(A) < 0 \quad \Leftrightarrow \quad \beta(A) > \epsilon \quad \Leftrightarrow \quad \|H\|_\infty < \frac{1}{\epsilon}. \tag{3.3}$$

We can characterize the condition $\alpha_\epsilon(A) < x$ analogously in terms of a "shifted" $\mathbf{H}_\infty$ norm [9, p. 67].

An important topic in robust control has been the design of controllers which minimize the $\mathbf{H}_\infty$ norm [44, 43]. In the language above, this corresponds to choosing the parameters defining the stable matrix A in order to maximize the minimum value of $\sigma_{\min}(A - zI)$ as z varies along the imaginary axis. Our ultimate aim of optimizing the pseudospectral abscissa is related, but rather different, being motivated by the broad idea of robust optimization [4]. We first fix the "level of robustness" ϵ (precisely the quantity that we seek to maximize in an $\mathbf{H}_\infty$ norm problem) and then vary A to move the corresponding pseudospectrum as far as possible to the left in the complex plane. In other words, we try to maximize a real parameter x such that the $\mathbf{H}_\infty$ norm corresponding to the shifted matrix $A - xI$ is not more than ϵ^{-1}.

What are the relative merits of different choices of ϵ in a pseudospectral minimization problem $\inf_F \alpha_\epsilon$? Here we are motivated by Trefethen's well-known viewpoint [39, 40], but we add an optimization "twist." When $\epsilon = 0$, optimization amounts to minimizing the spectral abscissa of a matrix $A \in F$, in other words, optimizing the *asymptotic* rate of decay of trajectories of the dynamical system $\dot{p} = Ap$. On the other hand, for large ϵ, by Theorem 3.2 (large ϵ), optimization amounts to minimizing $\lambda_{\max}(A + A^*)/2$. This corresponds to optimizing the *initial* decay rate of the dynamical system, since at time $t = 0$,

$$\frac{d}{dt}\frac{\|p\|^2}{2} = p(0)^*\Big(\frac{A + A^*}{2}\Big)p(0) \le \lambda_{\max}\Big(\frac{A + A^*}{2}\Big)\|p(0)\|^2,$$

with equality if $p(0)$ is an eigenvector corresponding to the largest eigenvalue. For intermediate choices of ϵ, minimizing the pseudospectral abscissa balances the two objectives of improving asymptotic stability and restricting the size of *transient* peaks in the trajectories. In particular, Proposition 3.1 (maximizing the distance to instability) shows that, under reasonable conditions, for some choice of ϵ, minimizing α_ϵ is equivalent to minimizing the $\mathbf{H}_\infty$ norm, that is, maximizing the complex stability radius.

To summarize, minimizing the $\mathbf{H}_\infty$ norm of a matrix A optimizes the robustness of the stability of the dynamical system $\dot{p} = Ap$, but with no explicit reference to its asymptotic decay rate. By minimizing the pseudospectral abscissa α_ϵ instead, for different choices of the parameter ϵ we obtain a range of different balances between robustness and asymptotic decay, one choice giving exactly the $\mathbf{H}_\infty$ norm problem. One could achieve a similar range of balances by minimizing the $\mathbf{H}_\infty$ norm corresponding to the shifted matrix $A - xI$ as the real parameter x varies; however, working with ϵ-pseudospectra for fixed ϵ provides a natural interpretation in terms of allowable perturbations to A. Yet another range of balances is achieved by the "robust spectral abscissa" defined in [14].

Just as with the $\mathbf{H}_\infty$ norm, the pseudospectral abscissa can be characterized via semidefinite programming. Specifically, by [9, p. 67] or [4, Prop. 4.4.2], a real x satisfies

$$\alpha_\epsilon(A) < x$$

if and only if there exist reals $\mu < 0$ and λ, and an $n \times n$ positive definite Hermitian matrix P such that the matrix

$$\begin{bmatrix} (\mu-\lambda)I + 2xP - A^*P - PA & -\epsilon P \\ -\epsilon P & \lambda I \end{bmatrix}$$

is positive semidefinite. As discussed in [9, pp. 3–4], the power of such semidefinite characterizations derives from their amenability to efficient interior point methods for convex optimization, pioneered in [31]. The disadvantage is the appearance of subsidiary semidefinite matrix variables: if the underlying matrices A are large, and we need to calculate the pseudospectral abscissa for many different matrices (in an optimization routine, for example), involving these subsidiary variables may be prohibitive computationally; see, for example, [33, 14, 42]. For this reason, in this work we consider more direct approaches to the pseudospectral abscissa.

4. Boundary properties. We begin our direct, geometric approach to the pseudospectral abscissa by studying the boundary of the pseudospectrum.

PROPOSITION 4.1 (compactness). *The pseudospectrum Λ_ϵ is a compact set contained in the ball of radius $\|A\| + \epsilon$. It contains the strict pseudospectrum Λ'_ϵ, which is nonempty and open.*

Proof. The strict pseudospectrum is nonempty since it contains the spectrum. It is open since $\sigma_{\min}$, and hence g are continuous. This also shows that the pseudospectrum is closed. For any point $z \in \Lambda_\epsilon$ there is a unit vector $u \in \mathbf{C}^n$ satisfying $\|(A-zI)u\| \le \epsilon$. On the other hand, $\|Au\| \le \|A\|$, so we have the inequality

$$|z| = \|zu\| \le \|(A-zI)u\| + \|Au\| \le \|A\| + \epsilon, \tag{4.1}$$

which shows boundedness. $\square$

The next result is slightly less immediate.

THEOREM 4.2 (local minima). *The only local minimizers of the function*

$$g(z) = \sigma_{\min}(A - zI)$$

are the eigenvalues of the matrix A.

Proof. Suppose the point z_0 is a local minimizer that is not an eigenvalue. Then z_0 is a local maximizer of the norm of the resolvent $\|(A-zI)^{-1}\|$. We can choose unit vectors $u, v \in \mathbf{C}^n$ satisfying

$$\|(A - z_0 I)^{-1}\| = |u^*(A - z_0 I)^{-1} v|,$$

and then we have, for all points z close to z_0, the inequalities

$$|u^*(A-zI)^{-1}v| \le \|(A-zI)^{-1}\| \le \|(A-z_0I)^{-1}\| = |u^*(A-z_0I)^{-1}v|.$$

Hence the modulus of the function $u^*(A-zI)^{-1}v$ has a local maximum at z_0. But this contradicts the maximum modulus principle, since this function is analytic and nonconstant near z_0. $\square$

COROLLARY 4.3 (closure of strict pseudospectrum). *The closure of the strict pseudospectrum is the pseudospectrum, so for $\epsilon > 0$ the pseudospectral abscissa is*

$$\alpha_\epsilon = \sup\{\operatorname{Re} z : z \in \Lambda'_\epsilon\}.$$

Proof. A point in the pseudospectrum that is outside the closure of the strict pseudospectrum must be a local minimizer of the function g. $\square$

An easy exercise now shows that the pseudospectral abscissa α_ϵ is a continuous, strictly increasing function of $\epsilon \in [0, +\infty)$. Note also that, by contrast with the above result, the function g may have local *maximizers* and, consequently, the strict pseudospectrum may not equal the interior of the pseudospectrum.

We can refine the above corollary with a more delicate argument, showing that we can "access" any point in the pseudospectrum via a smooth path through the strict pseudospectrum.

THEOREM 4.4 (accessibility). *Given any point z_0 in the pseudospectrum, there is a real-analytic path $p : [0,1] \to \mathbf{C}$ such that $p(0) = z_0$ and $p(t)$ lies in the strict pseudospectrum for all $t \in (0,1]$.*

Proof. We may as well assume $g(z_0) = \epsilon$. By Corollary 4.3, there exists a sequence $z_r \in \Lambda'_\epsilon$ approaching z_0. For each index r there exists a vector $u^r \in \mathbf{C}^n$ satisfying the inequalities

$$1 < \|u^r\| < 1 + \frac{1}{r} \quad \text{and} \quad \|(A - z_r I)u^r\| < \epsilon.$$

By taking a subsequence, we may as well assume that the sequence $\{u^r\}$ converges to a limit u^0, and then we have $(z_0, u^0) \in \operatorname{cl} S$, where

$$S = \Big\{(z,u) : \|u\|^2 > 1,\ \|(A - zI)u\|^2 < \epsilon^2\Big\}.$$

Since the set S is defined by a finite number of strict algebraic inequalities, we can apply the accessibility lemma [29]. Hence there is a real-analytic path $q : [0,1] \to \mathbf{C} \times \mathbf{C}^n$ such that $q(0) = (z_0, u^0)$ and $q(t) \in S$ for all $t \in (0,1]$. The result now follows by taking p to be the first component of q. □

In most cases the boundary of the pseudospectrum is straightforward to analyze without recourse to the above result. We make the following definition.

DEFINITION 4.5. *A point $z \in \mathbf{C}$ is* degenerate *if the smallest singular value of $A - zI$ is nonzero and simple (that is, has multiplicity one) and the corresponding right singular vector u satisfies $u^*(A - zI)u = 0$.*

We need the following elementary identity.

LEMMA 4.6. *Given any unit vector $u \in \mathbf{C}^n$, matrix $B \in \mathbf{M}^n$, and scalar $w \in \mathbf{C}$, we have*

$$\|(B + wI)u\|^2 - \|Bu\|^2 = |u^*(B + wI)u|^2 - |u^*Bu|^2.$$

The next result shows that, except possibly at degenerate points, the pseudospectrum can never be "pointed" outwards.

PROPOSITION 4.7 (pointedness). *Any nondegenerate point in the pseudospectrum lies on the boundary of an open disk contained in the strict pseudospectrum.*

Proof. Consider a nondegenerate point $z_0 \in \Lambda_\epsilon$. We may as well assume $g(z_0) = \epsilon$. Choose a unit right singular vector $u \in \mathbf{C}^n$ satisfying the condition $u^*(A - z_0 I)u \neq 0$. We now claim

$$|z - u^*Au| < |z_0 - u^*Au| \quad \Rightarrow \quad z \in \Lambda'_\epsilon.$$

To see this, observe that if z satisfies the left-hand side, then

$$\begin{aligned}
\sigma^2_{\min}(A - zI) - \epsilon^2 &\le \|(A - zI)u\|^2 - \|(A - z_0 I)u\|^2 \\
&= |u^*(A - zI)u|^2 - |u^*(A - z_0 I)u|^2 \\
&= |z - u^*Au|^2 - |z_0 - u^*Au|^2 \\
&< 0,
\end{aligned}$$

using the preceding lemma. ∎

In particular, this result shows that Theorem 4.4 is elementary in the case when the point of interest z_0 is nondegenerate.

Thus the pseudospectrum is not pointed outward, except possibly at a degenerate point. In fact, a more detailed analysis due to Trefethen shows that the pseudospectrum is *never* pointed outward [41]. However, it can certainly be pointed inward, as the following example shows.

Example 1 (nonsmooth points). Consider the matrix

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

The pseudospectrum Λ_ϵ consists of the union of two disks of radius ϵ, centered at the two eigenvalues, $\pm i$. For example, if $\epsilon = \sqrt{2}$, then $z = 1$ is a nonsmooth point where the boundary of the pseudospectrum is pointed inward. In the case where $\epsilon = 1$, the pseudospectrum consists of two disks tangent to each other at the origin.

Nonetheless, even though the boundary of the pseudospectrum can be nonsmooth, this cannot occur at any nondegenerate optimal solution of the pseudospectral abscissa problem.

PROPOSITION 4.8 (optimal solutions). *Any locally optimal solution z_0 of the pseudospectral abscissa problem* (2.1) *must lie on the boundary of the pseudospectrum. Furthermore, unless z_0 is degenerate, the boundary is differentiable there.*

Proof. The fact that z_0 cannot lie in the interior of Λ_ϵ is immediate. Now assume z_0 is nondegenerate. Since z_0 is optimal, Λ_ϵ lies on or to the left of the vertical line through z_0. But since z_0 is nondegenerate, Λ_ϵ contains a closed disk whose boundary contains z_0, by Proposition 4.7 (pointedness). Thus the boundary of Λ_ϵ lies between the disk and the vertical line, which are tangent at z_0. This completes the proof. ∎

Again, the nondegeneracy hypothesis may be dropped using the more general result on pointedness mentioned above [41].

5. Components of the pseudospectrum. We recall some basic ideas from plane topology. A *domain* is a nonempty, open, arcwise connected subset of $\mathbf{C}$. Given a point z in an open set $\Omega \subset \mathbf{C}$, a particular example of a domain is the *component* of z, which consists of all points that can be joined to z by a continuous path in Ω [35].

The following result is in essence well known (see, for example, [10]).

THEOREM 5.1 (eigenvalues and components). *Every component of the strict pseudospectrum of the matrix A contains an eigenvalue of A.*

Proof. Suppose the set S is a component of the strict pseudospectrum Λ'_ϵ that contains no eigenvalues of A. The function g attains its minimum on the compact set $\mathrm{cl}\, S$ at some point z, and clearly $g(z) < \epsilon$, so $z \in \Lambda'_\epsilon$. Since S is open and contains no eigenvalues, Theorem 4.2 (local minima) implies $z \notin S$.

But since Λ'_ϵ is open, it contains an open disk D centered at z. Since $z \in \mathrm{cl}\, S$, we know $D \cap S \neq \emptyset$, and hence $D \cup S$ is an arcwise connected subset of Λ'_ϵ strictly larger than S. But this contradicts the definition of S. ∎

In Example 1 (nonsmooth points), when $\epsilon = 1$ the strict pseudospectrum consists of two components, namely the two open disks centered at the two eigenvalues, $\pm i$. By contrast, the pseudospectrum is arcwise connected.

The simplest case of the above result occurs when each eigenvalue has geometric multiplicity one and ϵ is small. In this case we show later (Corollary 7.5) that the pseudospectrum consists of disjoint compact convex neighborhoods of each eigenvalue (cf. [32]).

Our next aim is to try to bracket the pseudospectral abscissa. We first need a subsidiary result.

LEMMA 5.2 (moving to the boundary). *For any point z_1 in Λ_ϵ there exists a point z_2 satisfying* $\operatorname{Re} z_1 = \operatorname{Re} z_2$ *and* $g(z_2) = \epsilon$.

Proof. We simply take z_2 on the boundary of the intersection of the vertical line through z_1 and the pseudospectrum Λ_ϵ (which is compact). ∎

Byers' algorithm for calculating the distance to instability [16] and its subsequent variants (see the introduction) all depend on versions of the following easy piece of linear algebra, relating singular values to imaginary eigenvalues of a certain Hamiltonian matrix. We include a proof for completeness.

LEMMA 5.3 (imaginary eigenvalues). *For real numbers x and y, and $\epsilon \geq 0$, the matrix $A - (x+iy)I$ has a singular value ϵ if and only if the matrix*

$$\begin{bmatrix} xI - A^* & \epsilon I \\ -\epsilon I & A - xI \end{bmatrix}$$

has an eigenvalue iy.

Proof. Plus and minus the singular values of any matrix $B \in \mathbf{M}^n$ are exactly the eigenvalues of the matrix

$$\begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix}.$$

Thus the matrix $A - (x+iy)I$ has a singular value ϵ if and only if ϵ is an eigenvalue of the matrix

$$\begin{bmatrix} 0 & A - (x+iy)I \\ A^* - (x-iy)I & 0 \end{bmatrix}$$

or, in other words, if and only if the matrix

$$\begin{bmatrix} -\epsilon I & A - (x+iy)I \\ A^* - (x-iy)I & -\epsilon I \end{bmatrix}$$

is singular. Since

$$\begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} -\epsilon I & A - (x+iy)I \\ A^* - (x-iy)I & -\epsilon I \end{bmatrix}$$
$$= \begin{bmatrix} (A^* - xI) & -\epsilon I \\ \epsilon I & (xI - A) \end{bmatrix} + iy \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix},$$

this is equivalent to iy being an eigenvalue of the matrix

$$\begin{bmatrix} xI - A^* & \epsilon I \\ -\epsilon I & A - xI \end{bmatrix}. \qquad ∎$$

The following result is our key test. Geometrically it states simply that a given real x (bigger than the spectral abscissa α) is less than the pseudospectral abscissa exactly when the vertical line through x intersects the boundary of the pseudospectrum. As we shall see, this is a straightforward computational test.

THEOREM 5.4 (bracketing the pseudospectral abscissa). *For any real $x \geq \alpha$, the following statements are equivalent:*

(i) $x \leq \alpha_\epsilon$;

(ii) *the equation*

$$g(x+iy)=\epsilon, \quad y\in\mathbf{R}, \tag{5.1}$$

is solvable;

(iii) *the system*

$$iy \in \Lambda \begin{bmatrix} xI - A^* & \epsilon I \\ -\epsilon I & A - xI \end{bmatrix}, \quad y \in \mathbf{R}, \tag{5.2}$$

is solvable.

Proof. We first show (i) $\Rightarrow$ (ii). If $x=\alpha_\epsilon$, then choose any point z solving the pseudospectral abscissa problem (2.1). Clearly $z=x+iy$ for some real y, and $g(z)=\epsilon$, so we have shown that (5.1) has a solution.

We can therefore assume $x<\alpha_\epsilon$, in which case there exists a point z_1 such that $\operatorname{Re} z_1 > x$ and $g(z_1)<\epsilon$. The component of z_1 in the strict pseudospectral abscissa Λ'_ϵ contains an eigenvalue z_2 by Theorem 5.1 (eigenvalues and components). Hence there is an arc in this component connecting z_1 and z_2. But since $\operatorname{Re} z_1 > x \geq \operatorname{Re} z_2$, this arc must contain a point z_3 with $\operatorname{Re} z_3 = x$. Now applying Lemma 5.2 (moving to the boundary) gives a solution to (5.1).

The implication (ii) $\Rightarrow$ (iii) is immediate from Lemma 5.3 (imaginary eigenvalues), so it remains to show (iii) $\Rightarrow$ (i). But this is again an easy consequence of Lemma 5.3: if system (5.2) holds, then ϵ is a singular value of the matrix $A-(x+iy)I$, and hence the smallest singular value of this matrix is no greater than ϵ, whence we get the result. □

Using this result, the relationship (3.3) between the pseudospectral abscissa and the $\mathbf{H}_\infty$ norm is an easy exercise.

We can now approximate the pseudospectral abscissa α_ϵ by a bisection search as follows.

ALGORITHM 5.5 (bisection method). We begin with the initial interval

$$[\alpha, \|A\|+\epsilon].$$

We know α_ϵ lies in this interval by the argument of Proposition 4.1 (compactness). Now at each iteration we let x be the midpoint of the current interval and compute all the eigenvalues of the matrix

$$\begin{bmatrix} xI - A^* & \epsilon I \\ -\epsilon I & A - xI \end{bmatrix}. \tag{5.3}$$

If any of the eigenvalues are purely imaginary, then we deduce $x \leq \alpha_\epsilon$ and replace the current interval with its right half. Otherwise, by Theorem 5.4 (bracketing the pseudospectral abscissa), we know $x>\alpha_\epsilon$, so we replace the current interval with its left half. The intervals generated by this algorithm are guaranteed to converge to the pseudospectral abscissa α_ϵ.

The difference between this algorithm and Byers' bisection method for the distance to instability [16] is that the former searches for x by bisection, while the latter searches for ϵ by bisection.

Notice that at each iteration of the bisection method we can easily solve (5.1). We first list the purely imaginary eigenvalues of the matrix (5.5), namely $\{iy_1, iy_2, \ldots, iy_k\}$. We then form the index set

$$J=\{j : \sigma_{\min}(A-(x+iy_j)I)=\epsilon\}.$$

The set of solutions of (5.1) is then simply $\{y_j : j \in J\}$. As we shall see in the next section, the points $x+iy_j$ (for $j \in J$) provide good approximations to *all* the solutions of the pseudospectral abscissa problem (2.1).

A more sophisticated, quadratically convergent algorithm for the pseudospectral abscissa, based on similar ideas and analogous to $\mathbf{H}_\infty$ norm algorithms such as [7, 11, 24], is developed in [13].

6. Approximate solutions. The results generated by the bisection algorithm (or the algorithm in [13]) approximate all the global maximizers in the pseudospectral abscissa problem (2.1). To make this precise we use the following standard notion of set convergence [34]. We say that a sequence of sets $Y^1, Y^2, \ldots \subset \mathbf{R}$ *converges* to a set $Y \subset \mathbf{R}$ if the following properties hold:

(i) For any number $y \in Y$ there exists a sequence of numbers $y^r \in Y^r$ converging to y;

(ii) any cluster point of a sequence of numbers $y^r \in Y^r$ lies in Y.

(This notion is weaker than the idea of convergence with respect to the Pompeiu–Hausdorff distance [34, Ex 4.13], although it is equivalent in the case when the sets Y^r and Y are uniformly bounded, as will be the case in our application below.)

We now prove a rather general result.

Theorem 6.1 (global maximizers). *The number of global maximizers of the pseudospectral abscissa problem* (2.1) *does not exceed n. Denote these*

$$\{\alpha_\epsilon + iy : y \in Y\},$$

where $Y \subset \mathbf{R}$. Consider any real sequence $\alpha \le x^r \uparrow \alpha_\epsilon$. Then the sets

$$Y_r = \{y \in \mathbf{R} : g(x^r + iy) = \epsilon\}$$

converge to Y.

Proof. The pseudospectral abscissa problem (2.1) has at least one maximizer, by compactness. Furthermore, any solution $z = \alpha_\epsilon + iy$ must satisfy the equation $g(z) = \epsilon$. Just as in the proof of Theorem 5.4 (bracketing the pseudospectral abscissa), this implies that y must satisfy the equation

$$\det \begin{bmatrix} -\epsilon I & A - (\alpha_\epsilon + iy)I \\ A^* - (\alpha_\epsilon - iy)I & -\epsilon I \end{bmatrix} = 0.$$

But this polynomial equation has at most $2n$ solutions, so we can write

$$Y = \{y_1, y_2, \ldots, y_m\},$$

where $1 \le m \le 2n$.

Fix the index $j \in \{1, 2, \ldots, m\}$. Theorem 4.4 (accessibility) and Theorem 5.1 (eigenvalues and components) together imply the existence of a continuous function $p : [0,1] \to \mathbf{C}$ such that $p(0) = \alpha_\epsilon + iy_j$, $p(1)$ is an eigenvalue, and $p(t) \in \Lambda'_\epsilon$ for all $t > 0$.

Using the continuity of p, we can now iteratively construct a nonincreasing sequence $\{t_r\} \subset [0,1]$ such that $\operatorname{Re} p(t_r) = x^r$ for all r. Taking limits shows

$$\operatorname{Re} p(\lim_r t_r) = \lim_r x^r = \alpha_\epsilon.$$

But for $t > 0$ we have $g(p(t)) < \epsilon$, which implies $\operatorname{Re} p(t) < \alpha_\epsilon$, so we deduce $t_r \downarrow 0$. Hence if we define $v_j^r = \operatorname{Im} p(t_r)$, we have $v_j^r \to y_j$.

For each index r, consider the bounded open set

$$\{y \in \mathbf{R} : g(x^r + iy) < \epsilon\}.$$

If $t_r = 0$, this set is empty, and we define $l_j^r = u_j^r = y_j$. Otherwise, denote the component of v_j^r in this set by the open interval (l_j^r, u_j^r). By continuity, l_j^r and u_j^r are both zeros of the function

$$y \in \mathbf{R} \mapsto g(x^r + iy) - \epsilon. \tag{6.1}$$

We now claim both $l_j^r \to y_j$ and $u_j^r \to y_j$.

If this claim fails, then without loss of generality, after taking a subsequence, we can assume $l_j^r \to w < y_j$. By definition, we know

$$g(x^r + i(sv_j^r + (1-s)l_j^r)) < \epsilon \ \text{ for all } s \in (0,1], \qquad r = 1, 2, \dots,$$

so taking limits shows

$$g(\alpha_\epsilon + i(sy_j + (1-s)w)) \le \epsilon \ \text{ for all } s \in [0,1].$$

But in this case every point in the line segment $\alpha_\epsilon + i[w, y_j]$ solves the pseudospectral abscissa problem (2.1), contradicting the fact that there are only finitely many solutions. This proves the claim. We have thus shown property (i) in the definition of set convergence: the constructed sequence (l_j^r) converges to the desired point y_j. Property (ii) is immediate.

Finally, suppose $m > n$. Choose any nondecreasing sequence $\{x^r\} \subset [\alpha, \alpha_\epsilon)$ converging to α_ϵ, and for each index r construct the set

$$\{l_j^r, u_j^r : j = 1, 2, \dots, m\}$$

as above. Then for r sufficiently large, this is a set of $2m$ distinct zeros of the function (6.1), and hence of the polynomial

$$\det \begin{bmatrix} -\epsilon I & A - (x^r + iy)I \\ A^* - (x^r - iy)I & -\epsilon I \end{bmatrix}.$$

But this polynomial is not identically zero, and has degree $2n$, which is a contradiction. $\square$

The algorithmic significance of the above result is this: Consider any algorithm that generates a sequence of lower approximations to the pseudospectral abscissa, $x^r \uparrow \alpha_\epsilon$. In particular, we could consider the bisection algorithm of the previous section. For each step r, an eigenvalue computation generates the set $Y^r \subset \mathbf{R}$, as described after Algorithm 5.5. The above result now shows that this set is a good approximation to Y, and hence gives us a good approximation to the set of all optimal solutions to the pseudospectral abscissa problem.

7. Smoothness. To study the smoothness of the function g, and hence the boundary of the pseudospectrum, we rely on the following well-known result. We consider $\mathbf{M}^n$ as a Euclidean space with inner product

$$\langle X, Y \rangle = \operatorname{Re} \operatorname{tr}(X^*Y) \quad (X, Y \in \mathbf{M}^n).$$

A real-valued function on a real vector space is *real-analytic* at zero if in some neighborhood of zero is can be written as the sum of an absolutely convergent power series in the coordinates relative to some basis, and we make an analogous definition at other points. In particular, such functions are smooth (C^∞) near the point in question.

We call vectors $u, v \in \mathbf{C}^n$ *minimal left and right singular vectors* for a matrix $Z \in \mathbf{M}^n$ if

$$Zv = \sigma_{\min}(Z)u \quad \text{and} \quad Z^*u = \sigma_{\min}(Z)v.$$

THEOREM 7.1 (analytic singular value). *If the matrix Z has a simple smallest singular value, then the function $\sigma^2_{\min}$ is real-analytic at Z. If, furthermore, $\sigma_{\min}(Z) > 0$, then $\sigma_{\min}$ is real-analytic at Z, with gradient*

$$\nabla\sigma_{\min}(Z) = uv^*$$

for any unit minimal left and right singular vectors $u, v \in \mathbf{C}^n$.

Proof. The matrix

$$(X^T - iY^T)(X + iY)$$

depends analytically on the matrices $X, Y \in \mathbf{M}^n$ and has a simple eigenvalue $\sigma^2_{\min}(Z)$ when $(X, Y) = (X_0, Y_0)$ for real matrices $X_0, Y_0 \in \mathbf{M}^n$ satisfying $Z = X_0 + iY_0$. Hence by standard perturbation theory [23], the above matrix has a unique eigenvalue near $\sigma^2_{\min}(Z)$ for all (X, Y) close to (X_0, Y_0), depending analytically on (X, Y). When X and Y are real, this eigenvalue is exactly $\sigma^2_{\min}(X + iY)$, so the first part follows. The second part follows by taking square roots. The gradient calculation is standard (see, for example, [37]). ∎

We next turn to smoothness properties of the function $g : \mathbf{C} \to \mathbf{R}$ defined by

$$g(z) = \sigma_{\min}(A - zI).$$

We will often find it more convenient to work with the squared function $g^2(z) = (g(z))^2$.

We can treat $\mathbf{C}$ as a Euclidean space, where we define the inner product by $\langle w, z\rangle = \mathrm{Re}\,(w^*z)$.

COROLLARY 7.2 (analytic boundary). *If the singular value $\sigma_{\min}(A - z_0I)$ is simple, then the function g^2 is real-analytic at z_0. If, furthermore, this singular value is strictly positive, then g is real-analytic at z_0, with gradient*

$$\nabla g(z_0) = -v^*u,$$

where the vectors $u, v \in \mathbf{C}^n$ are unit minimal left and right singular vectors for $A - z_0I$.

Proof. This follows from the previous result by the chain rule. ∎

Thus what we called "degenerate" points are simply smooth critical points of g, distinct from the eigenvalues. At a nondegenerate smooth point z_0 with $g(z_0) = \epsilon$, the gradient of g is nonzero, and hence the boundary of the pseudospectrum

$$\Lambda_\epsilon = \{z \in \mathbf{C} : g(z) \le \epsilon\}$$

is simply a smooth curve locally, with normal $u^*(A - z_0)u$ at z_0.

We call an eigenvalue of A *nonderogatory* if it has geometric multiplicity one. This is the most common type of multiple eigenvalue (from the perspective of the dimensions of the corresponding manifolds in $\mathbf{M}^n$ [1]). The following result is very well known.

PROPOSITION 7.3 (nonderogatory eigenvalues). *The point $\lambda_0 \in \mathbf{C}$ is a nonderogatory eigenvalue of the matrix A if and only if 0 is a simple singular value of $A - \lambda_0 I$.*

Proof. First, note that λ_0 is an eigenvalue of A if and only if $A - \lambda_0 I$ is singular, which is equivalent to 0 being a singular value of $A - \lambda_0 I$. Second, v is a corresponding eigenvector of A exactly when $(A - \lambda_0 I)v = 0$, which says that v is a right singular vector of $A - \lambda_0 I$ corresponding to the singular value 0. Thus the eigenspace of A corresponding to the eigenvalue λ_0 coincides with the subspace of right singular vectors of $A - \lambda_0 I$ corresponding to the singular value 0, so in particular these spaces have the same dimension. The result now follows. ∎

We can now show that the function g is well behaved near any nonderogatory eigenvalue of A.

THEOREM 7.4 (growth near an eigenvalue). *Let λ_0 be a nonderogatory eigenvalue of multiplicity m for the matrix A. Then*

$$\sigma_{\min}(A - zI) = g(z) = \frac{\prod_{j=1}^{n-m} |\lambda_j - \lambda_0|}{\prod_{k=1}^{n-1} \sigma_k} |z - \lambda_0|^m + O(|z - \lambda_0|^{m+1})$$

for complex z near λ_0, where $\lambda_1, \lambda_2, \ldots, \lambda_{n-m}$ are the eigenvalues of A distinct from λ_0 (listed by multiplicity) and $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}$ are the nonzero singular values of $A - \lambda_0 I$ (listed by multiplicity). (In the case $n = 1$ or $m = n$, we interpret the empty products appearing in the above expression as 1.)

Furthermore, the function g^2 has positive definite Hessian at all points $z \neq \lambda_0$ near λ_0.

Proof. We prove the case $\lambda_0 = 0$: the general case follows by a simple transformation.

Since 0 is a nonderogatory eigenvalue of A, Proposition 7.3 (nonderogatory eigenvalues) shows 0 is a simple singular value of A. Hence by Corollary 7.2 (analytic boundary), the function $f : \mathbf{R}^2 \to \mathbf{R}$ defined by

$$f(x, y) = (g(x + iy))^2$$

is real-analytic at $(0, 0)$.

Consider any point $(x, y) \in \mathbf{R}^2$, and let $z = x + iy$. The matrix

$$(A - zI)^*(A - zI)$$

is Hermitian, so its characteristic polynomial

$$p_z(\mu) = \det((A - zI)^*(A - zI) - \mu I)$$

has all real coefficients. Hence we can write

$$p_z(\mu) = \sum_{r=0}^{n} q_r(x, y)\mu^r$$

for some real polynomials q_r. The smallest zero of p_z is $f(x, y)$.

We concentrate on the two lowest-order coefficients of the above polynomial. First, note

$$q_0(x, y) = p_z(0)$$

$$
\begin{aligned}
&= \det((A - zI)^*(A - zI)) \\
&= |\det(A - zI)|^2 \\
&= |z|^{2m} \prod_{j=1}^{n-m} |\lambda_j - z|^2.
\end{aligned}
$$

Hence for small (x, y) we have

$$
q_0(x, y) = \left((x^2 + y^2)^m \prod_{j=1}^{n-m} |\lambda_j|^2 \right) + O(\|(x, y)\|^{2m+1}). \tag{7.1}
$$

Turning to the coefficient of μ, notice

$$
p_0(\mu) = \det(A^*A - \mu I) = -\mu \prod_{k=1}^{n-1} (\sigma_k^2 - \mu),
$$

so

$$
q_1(0, 0) = -\prod_{k=1}^{n-1} \sigma_k^2
$$

(notice this is nonzero), and hence

$$
q_1(x, y) = -\prod_{k=1}^{n-1} \sigma_k^2 + O(\|x, y\|). \tag{7.2}
$$

Since the function f is real-analytic at $(0, 0)$, we know for some integer $t = 1, 2, \ldots,$

$$
f(x, y) = s(x, y) + O(\|(x, y)\|^{t+1})
$$

for some nonzero homogeneous polynomial s of degree t. Now substituting into the relationship

$$
\sum_{r=0}^{n} q_r(x, y)(f(x, y))^r = 0
$$

and using (7.1) and (7.2) shows $t = 2m$, and

$$
f(x, y) = \frac{\prod_{j=1}^{n-m} |\lambda_j|^2}{\prod_{k=1}^{n-1} \sigma_k^2} (x^2 + y^2)^m + O(\|(x, y)\|^{2m+1})
$$

as required.

It remains to show that the Hessian $\nabla^2 f(x, y)$ is positive definite for all small $(x, y) \neq (0, 0)$. We have shown that f is analytic at $(0, 0)$ and

$$
f(x, y) = \tau(x^2 + y^2)^m + O(\|(x, y)\|^{2m+1})
$$

for some nonzero constant τ. Since we can differentiate the power series for f term-by-term, a short calculation shows

$$
\nabla^2 f(x, y) = 2m\tau(x^2 + y^2)^{m-2} H(x, y),
$$

where

$$H(x,y) = \begin{bmatrix} (2m-1)x^2 + y^2 & 2(m-1)xy \\ 2(m-1)xy & x^2 + (2m-1)y^2 \end{bmatrix} + O(\|(x,y)\|^3).$$

We deduce

$$H_{11}(x,y) = (2m-1)x^2 + y^2 + O(\|(x,y)\|^3) > 0$$

and furthermore

$$\det H(x,y) = (2m-1)(x^2+y^2)^2 + O(\|(x,y)\|^5) > 0,$$

so the matrix $H(x,y)$ is positive definite for all small $(x,y) \neq (0,0)$. The result now follows. ∎

The following results are immediate consequences.

COROLLARY 7.5 (convexity). *If λ_0 is a nonderogatory eigenvalue of the matrix A, then for all small $\epsilon > 0$ the pseudospectrum Λ_ϵ near λ_0 consists of a compact, strictly convex neighborhood of λ_0.*

Proof. We can consider the pseudospectrum as a level set of the function g^2, which is strictly convex near λ_0. ∎

COROLLARY 7.6 (smoothness). *If λ_0 is a nonderogatory eigenvalue of the matrix A, then the function g is smooth with nonzero gradient at all nearby points distinct from λ_0.*

Proof. Since the real-analytic function g^2 is strictly convex near the eigenvalue λ_0, with a strict local minimizer there, it follows that λ_0 is an isolated critical point of g^2. It is then easy to see that g is smooth and noncritical near λ_0. ∎

If the matrix A has an eigenvalue λ_0 of multiplicity m, then, by continuity of the set of eigenvalues, any matrix close to A will have exactly m eigenvalues close to λ_0 (counted by multiplicity). Our last corollary bounds how far these eigenvalues can be from λ_0.

COROLLARY 7.7 (Hölder continuity). *With the assumptions and notation of Theorem* 7.4, *consider any constant*

$$\kappa > \left(\frac{\prod_{k=1}^{n-1} \sigma_k}{\prod_{j=1}^{n-m} |\lambda_j - \lambda_0|} \right)^{1/m}.$$

For any matrix Z close to A, any eigenvalue z of Z close to λ_0 satisfies

$$|z - \lambda_0| \leq \kappa \|Z - A\|^{1/m}.$$

Proof. This follows easily from Theorem 7.4 (growth near an eigenvalue), using the elementary property that

$$\sigma_{\min}(A - zI) \leq \|Z - A\|$$

for any eigenvalue z of Z. ∎

In the above result, if we specialize to the case of a perturbation $Z = A + tB$ (where t is a complex parameter), then the result shows that the eigenvalues of $A + tB$ near a nonderogatory eigenvalue of A of multiplicity m satisfy an m^{-1}-Hölder continuity condition in t. This is a well-known result; see [27, 23, 3, 30].

8. Smoothness and regularity of the pseudospectral abscissa. Our ultimate goal is an understanding of how the pseudospectral abscissa α_ϵ depends on the underlying matrix A. We therefore now allow A (and ϵ) to vary. Recall that the pseudospectral abscissa function $\alpha_\epsilon : \mathbf{M}^n \to \mathbf{R}$ is given by

$$\alpha_\epsilon(Z) = \max\{\mathrm{Re}\, z : \sigma_{\min}(Z - zI) \le \epsilon\},$$

and for any nonempty set $\Omega \subset \mathbf{C}$ we define the refinement

$$\alpha^\Omega(Z, \epsilon) = \sup\{\mathrm{Re}\, z : z \in \Omega,\ \sigma_{\min}(Z - zI) \le \epsilon\}. \tag{8.1}$$

Thus for $\Omega = \mathbf{C}$, we obtain exactly the pseudospectral abscissa. We now apply classical sensitivity analysis to differentiate this function.

THEOREM 8.1 (smoothness of pseudospectral abscissa). *Suppose that, for $\epsilon = \epsilon_0 > 0$ and $Z = A$, the supremum* (8.1) *is attained by a point $z_0 \in \mathrm{int}\,\Omega$, where the singular value $\sigma_{\min}(A - z_0 I)$ is simple. Then for any corresponding unit minimal left and right singular vectors $u, v \in \mathbf{C}^n$, the number v^*u is real and nonpositive.*

Now suppose furthermore that z_0 is the unique attaining point in (8.1)*, that it is nondegenerate (or, in other words, $v^*u \ne 0$), and that the Hessian $\nabla^2(g^2)(z_0)$ is nonsingular. Then the function α^Ω is smooth around the point (A, ϵ_0), with*

$$\nabla_Z \alpha^\Omega(A, \epsilon_0) = \frac{uv^*}{v^*u} \quad \text{and} \quad \nabla_\epsilon \alpha^\Omega(A, \epsilon_0) = -\frac{1}{v^*u}.$$

Proof. Consider the optimization problem

$$\left\{\begin{array}{lrcl} \sup & \mathrm{Re}\, z & & \\ \text{subject to} & \sigma_{\min}^2(Z - zI) & \le & \epsilon^2, \\ & z & \in & \Omega. \end{array}\right.$$

When $(Z, \epsilon) = (A, \epsilon_0)$ this problem becomes

$$\left\{\begin{array}{lrcl} \sup & \mathrm{Re}\, z & & \\ \text{subject to} & g^2(z) & \le & \epsilon_0^2, \\ & z & \in & \Omega, \end{array}\right.$$

with optimal solution z_0. By Corollary 7.2 (analytic boundary), the function g^2 is smooth near z_0, with gradient

$$\nabla g^2(z_0) = 2g(z_0)\nabla g(z_0) = -2\epsilon_0 v^*u.$$

Either this gradient is zero or there is a Lagrange multiplier $\mu \in \mathbf{R}_+$ such that the gradient of the Lagrangian

$$z \mapsto \mathrm{Re}\, z - \mu(g^2(z) - \epsilon_0^2)$$

at $z = z_0$ is zero. In this case,

$$1 + 2\mu\epsilon_0 v^*u = 0, \tag{8.2}$$

so the first part follows.

Moving to the second part, (8.2) implies $\mu = -(2\epsilon_0 v^*u)^{-1}$. Under the additional assumptions we can apply a standard sensitivity result (for example, [6, Thm 5.5.3]) to deduce that the gradient of the optimal value of the original optimization problem at (A, ϵ_0) equals the gradient of the Lagrangian

$$(Z, \epsilon) \mapsto \operatorname{Re} z_0 + (2\epsilon_0 v^* u)^{-1}(\sigma^2_{\min}(Z - z_0 I) - \epsilon^2)$$

at (A, ϵ_0). The result now follows by Theorem 7.1 (analytic singular value). ∎

An eigenvalue of the matrix A with real part equal to the spectral abscissa α is called *active*.

THEOREM 8.2 (regular representation). *If the matrix A has s distinct active eigenvalues, all of which are nonderogatory, then there exist s functions*

$$\gamma_j : \mathbf{M}^n \times \mathbf{R}_{++} \to \mathbf{R} \quad (j = 1, 2, \ldots, s),$$

such that for small $\epsilon > 0$ and matrices Z close to A, each map

$$(Z, \epsilon) \mapsto \gamma_j(Z, \epsilon)$$

is smooth and satisfies $\gamma_j(A, 0) = \alpha(A)$, the pseudospectral abscissa can be expressed as

$$\alpha_\epsilon(Z) = \max\{\gamma_j(Z, \epsilon) : j = 1, 2, \ldots, s\},$$

and the set of gradients

$$\{\nabla_Z \gamma_j(A, \epsilon) : j = 1, 2, \ldots, s\}$$

is linearly independent.

Proof. Denote the distinct eigenvalues of A by $\lambda_1, \lambda_2, \ldots, \lambda_k$, where

$$\operatorname{Re} \lambda_j \begin{cases} = & \alpha \quad (j \leq s), \\ < & \alpha \quad (j > s). \end{cases}$$

Let D denote the open unit disk in $\mathbf{C}$. Providing we choose a radius $\delta > 0$ sufficiently small, we have

$$\begin{aligned} 2\delta &< |\lambda_p - \lambda_q| \ \text{ for all } p \neq q, \\ \delta + \operatorname{Re} \lambda_j &< \alpha \ \text{ for all } j > s, \end{aligned}$$

and so the open disks $\lambda_j + \delta D$ are disjoint, and those with $j > m$ lie in the half-plane $\operatorname{Re} z < \alpha$. Furthermore, again by reducing δ if necessary, Theorem 7.4 (growth near eigenvalues) guarantees that each of the functions

$$g^2\big|_{\lambda_j + \delta D} \quad (j = 1, 2, \ldots, s)$$

is smooth, with everywhere positive definite Hessian except possibly at λ_j.

We claim that the small pseudospectra (by which we mean pseudospectra corresponding to small ϵ) of matrices close to A lie in small disks around the eigenvalues of A. More precisely, for small $\epsilon \geq 0$ and matrices Z close to A, we claim

$$\{z \in \mathbf{C} : \sigma_{\min}(Z - zI) \leq \epsilon\} \subset \{\lambda_1, \lambda_2, \ldots, \lambda_k\} + \delta D. \tag{8.3}$$

Otherwise there would exist sequences $\epsilon_r \to 0$, $Z_r \to A$, and $z_r \in \mathbf{C}$ satisfying, for all $r = 1, 2, \ldots,$

$$\begin{aligned} \sigma_{\min}(Z_r - z_r I) &\leq \epsilon_r, \\ |z_r - \lambda_j| &\geq \delta \quad (j = 1, 2, \ldots, k). \end{aligned}$$

The first inequality above implies the sequence $\{z_r\}$ is bounded, so has a cluster point z_0, which must satisfy the inequalities

$$\begin{aligned}\sigma_{\min}(A - z_0 I) &\leq 0,\\ |z_0 - \lambda_j| &\geq \delta \quad (j = 1, 2, \ldots, k).\end{aligned}$$

The first inequality above can only hold if z_0 is an eigenvalue of A, which contradicts the second inequality. Hence inequality (8.3) holds, as we claimed.

Using the notation of (8.1), we can, for small $\epsilon > 0$, matrices Z close to A and, for each $j = 1, 2, \ldots, m$, define functions

$$\gamma_j(Z, \epsilon) = \alpha^{\lambda_j + \delta D}(Z, \epsilon) = \sup\{\mathrm{Re}\, z : |z - \lambda_j| < \delta,\ \sigma_{\min}(Z - zI) \leq \epsilon\},$$

and as a consequence of inclusion (8.3), we can then write

$$\alpha_\epsilon(Z) = \max\{\gamma_j(Z, \epsilon) : j = 1, 2, \ldots, m\}.$$

We claim each function γ_j is smooth around the point (A, ϵ_0) for any small $\epsilon_0 > 0$.

To prove this claim, we use Theorem 8.1 (smoothness of pseudospectral abscissa). For any $j = 1, 2, \ldots, s$, consider the supremum

$$\begin{aligned}\gamma_j(A, \epsilon_0) &= \sup\{\mathrm{Re}\, z : |z - \lambda_j| < \delta,\ \sigma_{\min}(A - zI) \leq \epsilon_0\}\\ &= \sup\left\{\mathrm{Re}\, z : \left. g^2\right|_{\lambda_j + \delta D}(z) \leq \epsilon_0^2\right\}.\end{aligned}$$

By our choice of the radius δ, this supremum is attained at a unique point z_j (cf. Corollary 7.5 (convexity)), which is nondegenerate (cf. Corollary 7.6 (smoothness)), and at which the Hessian $\nabla^2(g^2)(z_j)$ is positive definite. Hence the function γ_j is smooth around (A, ϵ_0), with gradient

$$\nabla_Z \gamma_j(A, \epsilon_0) = \frac{u_j v_j^*}{v_j^* u_j},$$

where u_j, v_j are unit minimal left and right singular vectors for $A - z_j I$, and $v_j^* u_j$ is real and strictly negative.

To complete the proof, it suffices to show that the set of matrices

$$\{u_j v_j^* : j = 1, 2, \ldots, s\} \subset \mathbf{M}^n$$

is linearly independent providing our choice of radius $\delta > 0$ is sufficiently small. If this fails, then for each j there is a sequence of points $z_j^r \to \lambda_j$ and sequences of unit minimal left and right singular vectors u_j^r, v_j^r for $A - z_j^r I$ such that the set of matrices

$$\{u_j^r (v_j^r)^* : j = 1, 2, \ldots, s\} \subset \mathbf{M}^n$$

is linearly dependent. By taking subsequences, we can suppose $u_j^r \to u_j^0$ and $v_j^r \to v_j^0$ for each j, and then the set

$$S = \{u_j^0 (v_j^0)^* : j = 1, 2, \ldots, s\} \subset \mathbf{M}^n$$

must be linearly dependent. But it also follows that u_j^0, v_j^0 are unit left and right eigenvectors for A corresponding to the eigenvalue λ_j. Since the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_s$ are distinct, the sets of eigenvectors

$$\{u_j^0 : j = 1, 2, \ldots, m\} \ \text{ and } \ \{v_j^0 : j = 1, 2, \ldots, s\}$$

are each linearly independent, and a standard exercise then shows the contradiction that the set S above is linearly independent. ∎

As a consequence of this result, the pseudospectral abscissa must be a reasonably well-behaved nonsmooth function near a matrix with all nonderogatory eigenvalues. Specifically, we have the following result. We refer the reader to [17, 34] for standard nonsmooth terminology.

COROLLARY 8.3 (regularity). *If all the active eigenvalues of a matrix $A \in \mathbf{M}^n$ are nonderogatory, then for all small $\epsilon > 0$ the pseudospectral abscissa α_ϵ is locally Lipschitz and subdifferentially regular around A.*

Proof. This follows immediately from the representation as a maximum of smooth functions in the previous result [17, Prop. 2.3.12]. ∎

This corollary presents an interesting parallel with a key result in [15]. This result states that the spectral abscissa, even though non-Lipschitz, is a subdifferentially regular function around the matrix A if and only if each active eigenvalue of A is nonderogatory.

By combining the representation of α_ϵ constructed in the proof of Theorem 8.2 with the growth estimate of Theorem 7.4, we can also see how the pseudospectral abscissa depends on the parameter ϵ.

COROLLARY 8.4 (dependence on ϵ). *If all the active eigenvalues of the matrix A are nonderogatory, with maximum algebraic multiplicity m, then as a function of $\epsilon \geq 0$ we have*

$$\alpha_\epsilon - \alpha \sim \gamma \epsilon^{1/m} \quad \textit{as } \epsilon \downarrow 0,$$

for some constant $\gamma > 0$.

9. Nonsmooth geometry. What about points $z_0 \in \mathbf{C}$, where $\sigma_{\min}(A - z_0 I)$ is multiple? The function $\sigma_{\min}$ is nonsmooth at any matrix with a multiple smallest singular value, so the function g may be nonsmooth at z_0. An appropriate approach to studying the pseudospectrum near z_0 is therefore to use nonsmooth analysis. We again refer to [17, 34] for the standard concepts.

For any point $z \in \mathbf{C}$ we consider the subspace $U(z) \subset \mathbf{C}^n$ spanned by all right singular vectors corresponding to $\sigma_{\min}(A - zI)$, and we define a subset of $\mathbf{C}$ by

$$G(z) = \{u^*(A - zI)u : u \in U(z),\ \|u\| = 1\}.$$

PROPOSITION 9.1 (convexity). *The set $G(z)$ is nonempty, compact, and convex.*

Proof. Define a linear map $B : U(z) \to U(z)$ by

$$Bu = P_{U(z)}((A - zI)u),$$

where $P_{U(z)} : \mathbf{C}^n \to U(z)$ denotes the orthogonal projection. Now notice for all vectors $u \in U(z)$ we have

$$\begin{aligned}\langle u, Bu\rangle = \langle u, P_{U(z)}((A - zI)u)\rangle &= \langle P^*_{U(z)} u, (A - zI)u\rangle \\ &= \langle u, (A - zI)u\rangle = u^*(A - zI)u,\end{aligned}$$

since the map $P^*_{U(z)} : U(z) \to \mathbf{C}^n$ is just the embedding. We deduce

$$G(z) = \{\langle u, Bu\rangle : u \in U(z),\ \|u\| = 1\},$$

and this set is nonempty, compact, and convex, by the Toeplitz–Hausdorff theorem [22]. ∎

The next result gives another perspective on the pointedness of the pseudospectrum (recall Proposition 4.7 (pointedness)).

THEOREM 9.2 (nonsmooth boundary behavior). *For complex z_0 satisfying $g(z_0) = \epsilon$ and $0 \notin G(z_0)$, the complement of the strict pseudospectrum,*

$$\{z \in \mathbf{C} : g(z) \geq \epsilon\},$$

is Clarke regular at z_0, with normal cone $\mathrm{cone}(G(z_0))$.

Proof. The complement of the strict pseudospectrum is

$$\begin{aligned} &\{z \in \mathbf{C} : \sigma_{\min}(A - zI) \geq \epsilon\} \\ &= \{z : \lambda_{\min}((A - zI)^*(A - zI)) \geq \epsilon^2\} \\ &= \{z : F(z) \in \mathbf{H}^n_+\} = F^{-1}(\mathbf{H}^n_+), \end{aligned}$$

where $\mathbf{H}^n$ denotes the Euclidean space of $n \times n$ Hermitian matrices, with inner product $\langle X, Y\rangle = \mathrm{Re}\,(\mathrm{tr}\,(XY))$ and positive semidefinite cone $\mathbf{H}^n_+$, the function $\lambda_{\min} : \mathbf{H}^n \to \mathbf{R}$ is the smallest eigenvalue, and the function $F : \mathbf{C} \to \mathbf{H}^n$ is defined by

$$F(z) = (A - zI)^*(A - zI) - \epsilon^2 I.$$

The gradient map $\nabla F(z_0) : \mathbf{C} \to \mathbf{H}^n$ is given by

$$\nabla F(z_0)(w) = -w^* A - wA^* + 2\langle w, z_0\rangle I,$$

and a short calculation shows that the adjoint map $\nabla F(z_0)^* : \mathbf{H}^n \to \mathbf{C}$ is given by

$$\nabla F(z_0)^* X = 2\mathrm{tr}\,((z_0 I - A)X).$$

It is well known (see, for example, [25]) that the positive semidefinite cone is Clarke regular at $F(z_0)$ (being convex), with normal cone

$$\begin{aligned} N_{\mathbf{H}^n_+}(F(z_0)) &= -\mathrm{cone}\{uu^* : F(z_0)u = 0\} \\ &= -\mathrm{cone}\{uu^* : u \in U(z_0),\ \|u\| = 1\}. \end{aligned}$$

Now consider any matrix

$$X \in N_{\mathbf{H}^n_+}(F(z_0)) \cap N(\nabla F(z_0)^*).$$

By the calculations above, we deduce

$$-X = \sum_{j=1}^{k} \mu_j u_j u_j^*$$

for some integer k, reals $\mu_j \geq 0$, and unit vectors $u_j \in U(z_0)$ $(j = 1, 2, \ldots, k)$, and

$$0 = \mathrm{tr}\,((A - z_0 I)X) = \sum_{j=1}^{k} \mu_j u_j^*(A - z_0 I)u_j.$$

But since $0 \notin G(z_0)$, by Proposition 9.1 (convexity) this implies that each μ_j is zero. We have therefore proved the condition

$$N_{\mathbf{H}^n_+}(F(z_0)) \cap N(\nabla F(z_0)^*) = \{0\}.$$

Under this condition we can apply a standard chain rule [34] to the set of interest, $F^{-1}(\mathbf{H}_+^n)$, to deduce that it is Clarke regular at the point z_0, with normal cone

$$N_{F^{-1}(\mathbf{H}_+^n)}(z_0) = \nabla F(z_0)^* N_{\mathbf{H}_+^n}(F(z_0)) = \mathrm{cone}(G(z_0)),$$

as required. □

Consider, for instance, Example 1 (nonsmooth points). When $\epsilon = \sqrt{2}$, we saw that the point $z_0 = 1$ is a nonsmooth point on the boundary of the pseudospectrum, which consists of the union of two disks of radius $\sqrt{2}$, centered at $\pm i$. A calculation shows that the set $G(z_0)$ in this case is the line segment $[1-i, 1+i]$, so according to the above result, the normal cone to the complement of the strict pseudospectrum is the cone $\{x + iy : |y| \leq -x\}$, as we expect.

By contrast, when $\epsilon = 1$ we saw that the pseudospectrum consists of two unit disks, tangent at 0. A calculation shows $G(0)$ is the line segment $[-i, i]$, which contains 0, so the above theorem does not apply.

Acknowledgments. Many thanks to Peter Benner, Carsten Scherer, and two anonymous referees for many helpful suggestions and references.

REFERENCES

[1] V. I. Arnold, *On matrices depending on parameters*, Russian Math. Surveys, 26 (1971), pp. 29–43.

[2] V. Balakrishnan, S. Boyd, and S. Balemi, *Branch and bound algorithm for computing the minimum stability degree of parameter-dependent linear systems*, Internat. J. Robust Nonlinear Control, 1 (1991), pp. 295–317.

[3] H. Baumgärtel, *Analytic Perturbation Theory for Matrices and Operators*, Birkhäuser-Verlag, Basel, 1985.

[4] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, SIAM, Philadelphia, 2001.

[5] V. Blondel and J. N. Tsitsiklis, *NP-hardness of some linear control design problems*, SIAM J. Control Optim., 35 (1997), pp. 2118–2127.

[6] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[7] S. Boyd and V. Balakrishnan, *A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞-norm*, Systems Control Lett., 15 (1990), pp. 1–7.

[8] S. Boyd, V. Balakrishnan, and P. Katamba, *A bisection method for computing the H_∞ norm of a transfer matrix and related problems*, Math. Control Signals Systems, 2 (1989), pp. 207–219.

[9] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, 1994.

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

[11] N. A. Bruinsma and M. Steinbuch, *A fast algorithm to compute the $\mathbf{H}_\infty$-norm of a transfer function matrix*, Systems Control Lett., 14 (1990), pp. 287–293.

[12] J. V. Burke, A. S. Lewis, and M. L. Overton, *Optimal stability and eigenvalue multiplicity*, Foundations Comput. Math., 1 (2001), pp. 205–225.

[13] J. V. Burke, A. S. Lewis, and M. L. Overton, *Robust stability and a criss-cross algorithm for pseudospectra*, IMA J. Numer. Anal., to appear.

[14] J. V. Burke, A. S. Lewis, and M. L. Overton, *Two numerical methods for optimizing matrix stability*, Linear Algebra Appl., 351/352 (2002), pp. 117–145.

[15] J. V. Burke and M. L. Overton, *Variational analysis of non-Lipschitz spectral functions*, Math. Programming, 90 (2001), pp. 317–352.

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

[17] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983. Republished as Classics in Applied Math. 5, SIAM, Philadelphia, 1990.

[18] Y. Genin, P. Van Dooren, and V. Vermaut, *Convergence of the calculation of* $\mathbf{H}_\infty$ *norms and related questions*, in Proceedings MTNS-98, G. Picci and D. S. Gilliam, eds., Birkhäuser-Verlag, Basel, 1998.
[19] C. He and G. A. Watson, *An algorithm for computing the distance to instability*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 101–116.
[20] *Pseudospectra Gateway*, http://web.comlab.ox.ac.uk/projects/pseudospectra.
[21] D. Hinrichsen and A. J. Pritchard, *Stability radii of linear systems*, Systems Control Lett., 7 (1986), pp. 1–10.
[22] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.
[23] T. Kato, *A Short Introduction to Perturbation Theory for Linear Operators*, Springer, New York, 1982.
[24] C. Lawrence, A. L. Tits, and P. Van Dooren, *A fast algorithm for the computation of an upper bound on the* μ*-norm*, Automatica, 36 (2000), pp. 449–456.
[25] A. S. Lewis, *Convex analysis on the Hermitian matrices*, SIAM J. Optim., 6 (1996), pp. 164–177.
[26] A. S. Lewis, *Robust regularization*, Math. Program., submitted.
[27] V. B. Lidskii, *Perturbation theory of non-conjugate operators*, USSR Comput. Math. and Math. Phys., 1 (1965), pp. 73–85.
[28] C. F. Van Loan, *How near is a stable matrix to an unstable matrix?* Contemp. Math., 47 (1985), pp. 465–477.
[29] J. W. Milnor, *Singular Points of Complex Hypersurfaces*, Princeton University Press, Princeton, NJ, 1968.
[30] J. Moro, J. V. Burke, and M. L. Overton, *On the Lidskii–Vishik–Lyusternik perturbation theory for the eigenvalues of matrices with arbitrary Jordan structure*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 793–817.
[31] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.
[32] B. T. Polyak, *Convexity of nonlinear image of a small ball with applications to optimization*, Set-Valued Anal., 9 (2001), pp. 159–168.
[33] X. Rao, K. Gallivan, and P. Van Dooren, *Stabilization of large scale dynamical systems*, in Proceedings MTNS 2000, S. L. Campbell, ed., Int. J. Appl. Math. Comput. Sci. II, University of Zielona Góra, Zielona, Góra, Poland, 2001.
[34] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
[35] R. A. Silverman, *Introductory Complex Analysis*, Prentice–Hall, Englewood Cliffs, NJ, 1967.
[36] J. Sreedhar, P. Van Dooren, and A. L. Tits, *A fast algorithm to compute the real structured stability radius*, in Stability Theory: Proceedings of Hurwitz Centenary Conference, R. Jeltsch and M. Mansour, eds., Birkhäuser-Verlag, Basel, 1996, pp. 219–230.
[37] G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
[38] M. J. Todd, *Semidefinite optimization*, Acta Numer., 10 (2001), pp. 515–560.
[39] L. N. Trefethen, *Pseudospectra of linear operators*, SIAM Rev., 39 (1997), pp. 383–406.
[40] L. N. Trefethen, *Computation of pseudospectra*, Acta Numer., 8 (1999), pp. 247–295.
[41] L. N. Trefethen, *private communication*, 2002.
[42] A. Varga and P. Parrilo, *Fast algorithms for solving* $\mathbf{H}_\infty$ *norm minimization problems*, in Proceedings of the 40th IEEE Conference on Decision and Control (Orlando, FL), IEEE, Piscataway, NJ, 2001.
[43] M. Vidyasagar, *Control Systems Synthesis: A Factorization Approach*, MIT Press, Cambridge, MA, 1985.
[44] G. Zames and B. A. Francis, *Feedback, minimax sensitivity, and optimal robustness*, IEEE Trans. Automat. Control, 28 (1983), pp. 585–601.