CONVEXITY AND LIPSCHITZ BEHAVIOR OF SMALL PSEUDOSPECTRA*

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 complex matrices within a distance ϵ of A, measured by the operator 2-norm. Given a nonderogatory matrix A_0, for small $\epsilon > 0$, we show that the ϵ-pseudospectrum of any matrix A near A_0 consists of compact convex neighborhoods of the eigenvalues of A_0. Furthermore, the dependence of each of these neighborhoods on A is Lipschitz.

Key words. pseudospectrum, eigenvalue optimization, nonsmooth analysis, Lipschitz multifunction, robust optimization

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

DOI. 10.1137/050645841

1. Introduction. Given a matrix A in the space of $n \times n$ complex matrices $\mathbf{M}^n$, the spectrum $\Lambda(A)$ is an informative analytic tool, but must be interpreted with care. In particular, when A has a multiple eigenvalue, small perturbations cause the spectrum to behave in a non-Lipschitz fashion.

Pseudospectra are robust analogs of the spectrum, enjoying many useful modelling properties. A comprehensive reference is [9]. We denote the operator 2-norm on $\mathbf{M}^n$ by $\| \cdot \|$. For real $\epsilon > 0$, the ϵ-*pseudospectrum* of A is the subset of the complex plane consisting of all eigenvalues of all complex matrices within a distance ϵ of A, measured by the operator 2-norm:

$$(1.1) \qquad \Lambda_\epsilon(A) = \bigcup_{\|X-A\| \le \epsilon} \Lambda(X).$$

This subset of the complex plane $\mathbf{C}$ is semialgebraic (meaning that it can be described as a finite union of sets each defined via finitely many polynomial inequalities [2]) and consists of at most n connected components; each component is compact and contains an eigenvalue of A. Visual plots of pseudospectra are richly informative and are conveniently computable via the EigTool package [5]. Note that, by contrast to our definition, [9] defines pseudospectra via the *strict* inequality $\|X - A\| < \epsilon$.

Our aim in this work is to show how shifting attention from the spectrum to pseudospectra has a regularizing effect on variational behavior. Specifically, for matrices A that are in a certain sense typical, even in the presence of multiple eigenvalues, if

the parameter ϵ is small, the ϵ-pseudospectrum consists of compact convex neighborhoods of the eigenvalues and varies in a Lipschitz fashion with respect to the Hausdorff distance.

2. Examples. We begin with two examples to illustrate the potential difficulties. We first observe how the pseudospectrum can vary in a non-Lipschitz fashion even around a two-by-two matrix with simple eigenvalues. Second, we note that the component of the ϵ-pseudospectrum containing a derogatory eigenvalue may fail to be convex, no matter how small the parameter $\epsilon > 0$.

To help our calculations, we use a well-known description of the pseudospectrum, more convenient than the definition (1.1). Denoting the smallest singular value by $\sigma_{\min} : \mathbf{M}^n \to \mathbf{R}$, the pseudospectrum is related to the reciprocal of the norm of the resolvent,

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

via the useful characterization

$$\Lambda_\epsilon(A) = \{z \in \mathbf{C} : \sigma_{\min}(A - zI) \leq \epsilon\}.$$

For our first example, we consider the behavior of the pseudospectrum $\Lambda_{\phi-1}(\cdot)$, where ϕ is the golden ratio $(1 + \sqrt{5})/2$, for matrices close to

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

An elementary calculation shows, for real r and θ, the formula

$$(2.1) \qquad 2\sigma_{\min}^2(\widehat{A} - re^{i\theta}I) \;=\; 3 + 2r^2 - \sqrt{5 + 4r^2(3 + 2\cos 2\theta)}.$$

This leads to a description of the pseudospectrum of $\widehat{A}$:

$$\Lambda_{\phi-1}(\widehat{A}) \;=\; \{re^{i\theta} : r^2 \leq 2(2 - \phi + \cos 2\theta)\}.$$

The boundary of this set is a lemniscate centered at zero (see Figure 2.1); its interior consists of two disjoint open sets, each containing one of the eigenvalues ± 1. In particular, notice that the pseudospectrum is contained in its tangent cone at zero:

$$(2.2) \qquad\qquad \Lambda_{\phi-1}(\widehat{A}) \;\subset\; \{re^{i\theta} : \cos 2\theta \geq \phi - 2\}.$$

Now consider the point ri on the imaginary axis as $r \downarrow 0$. The inclusion (2.2) implies a lower bound on the distance from this point to the pseudospectrum $\Lambda_{\phi-1}(\widehat{A})$ of the form
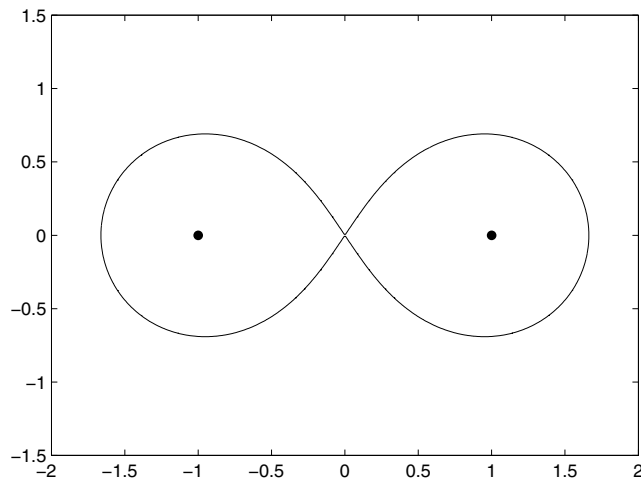
$$(2.3) \qquad\qquad d(ri, \Lambda_{\phi-1}(\widehat{A})) \geq \alpha r,$$

for some constant $\alpha > 0$. On the other hand, formula (2.1) implies

$$(2.4) \qquad\qquad \sigma_{\min}(\widehat{A} - re^{i\theta}I) = \phi - 1 + O(r^2),$$

so for some constant $\beta > 0$ we know that

$$ri \in \Lambda_{\phi-1+\beta r^2}(\widehat{A}).$$

FIG. 2.1. *Pseudospectrum of $\widehat{A}$.*

Using the definition of the pseudospectrum (1.1), we can rewrite the right-hand side as

$$\bigcup_{\|A-\widehat{A}\|\leq\beta r^2} \Lambda_{\phi-1}(A),$$

so there exists a matrix A_r satisfying

(2.5) $$\|A_r - \widehat{A}\| \leq \beta r^2 \quad \text{and} \quad ri \in \Lambda_{\phi-1}(A_r).$$

The *Hausdorff distance* between two nonempty sets $K, L \subset \mathbf{C}$ is the quantity

$$H(K,L) = \max\left\{ \sup_{z\in K} d(z,L), \sup_{z\in L} d(z,K) \right\},$$

where $d(z,L)$ is the distance from z to L. Now, in conjunction with inequality (2.3), the relationships (2.5) imply that the Hausdorff distance between the pseudospectra $\Lambda_{\phi-1}(\widehat{A})$ and $\Lambda_{\phi-1}(A_r)$ is at least αr, and yet the distance between the matrices $\widehat{A}$ and A_r is at most βr^2. Thus the variation of the mapping $\Lambda_{\phi-1}$ around $\widehat{A}$ is not Lipschitz.

The pathology in this example is caused by the existence of a critical point of the function $z \mapsto \sigma_{\min}(\widehat{A} - zI)$ at a point on the boundary of the pseudospectrum (in this case $z = 0$): this can be seen directly from formula (2.4), or by observing that the left and right singular vectors of the matrix $\widehat{A}$ corresponding to the smallest singular value $\phi - 1$ are orthogonal (see [3, Cor. 7.2]). A direct calculation is also illuminating. Since $\sigma_{\min}(\widehat{A}) = \phi - 1$, replacing by zero the diagonal entry $\phi - 1$ in the singular value decomposition of $\widehat{A}$ makes a perturbation of size $\phi - 1$ and results in a singular matrix. But a straightforward calculation shows that this singular matrix is similar to a two-by-two Jordan block, so further perturbations of size δ result in the zero eigenvalue splitting into two distinct eigenvalues of size proportional to $\sqrt{\delta}$. It

is this splitting that causes the pseudospectrum to behave in a non-Lipschitz fashion. In the development that follows, we avoid this possibility by focusing on the case of small ϵ.

We discuss various aspects of the growth of pseudospectra as the parameter ϵ grows in a forthcoming work [4]. In particular, we can quantitatively estimate the component of the pseudospectrum $\Lambda_\epsilon(A)$ containing the eigenvalue λ: classical eigenvalue perturbation theory shows that the component approximates a disk of radius $(\alpha\epsilon)^{1/m}$ as $\epsilon \downarrow 0$, where m is the multiplicity of λ as a root of the minimal polynomial for A, and α is its associated condition number [4].

Despite approximating disks, small pseudospectral components may be nonconvex in general, as shown by our second example, suggested by [8]. Consider the matrix

$$\tilde{A} = \left[\begin{array}{cc} 0 & 1 \\ 0 & 1 \end{array} \right].$$

An easy calculation shows

$$f(r, \theta) = \sigma_{\min}^2(\tilde{A} - re^{i\theta}I) = 1 - r\cos\theta + r^2 - \sqrt{(1 - r\cos\theta)^2 + r^2}$$

$$= \frac{r^2}{2}(1 - r\cos\theta) + O(r^4) \ \text{ as } r \downarrow 0.$$

Hence the component of the pseudospectrum $\Lambda_\epsilon(\tilde{A})$ containing zero, which we denote $\Lambda_\epsilon^0(\tilde{A})$, is a slightly distorted disk centered at zero and with radius approximately $\sqrt{2}\epsilon$, for small $\epsilon > 0$.

When $\theta = \pi/2$, another calculation shows

$$\frac{\partial f}{\partial r} = 2r\left(1 - \frac{1}{\sqrt{1 + r^2}}\right) > 0 \ \text{ for all } r > 0.$$

Hence for θ near $\pi/2$, the equation $f(r, \theta) = \epsilon^2$ implicitly defines r as a smooth function $g(\theta)$, and for r near $g(\pi/2) = \sqrt{2}\epsilon + O(\epsilon^2)$, the pseudospectrum is

$$\{re^{i\theta} : r \leq g(\theta)\}.$$

One more calculation shows

$$g'(\pi/2) = \frac{\sqrt{1 + g^2(\pi/2)} - 1}{2\sqrt{1 + g^2(\pi/2)} - 1} = \epsilon^2 + O(\epsilon^3).$$

To summarize, the pseudospectral boundary for the matrix $\tilde{A}$ crosses the positive imaginary axis at a unique point $z_\epsilon = (\sqrt{2}\epsilon + O(\epsilon^2))i$. The boundary nearby is a smooth curve crossing the imaginary axis nonorthogonally and bounding the pseudospectral component below it. Clearly, exactly the same properties hold for the matrix $-\tilde{A}$, and the two boundaries are mirror images in the imaginary axis. Finally, consider the matrix

$$A = \left[\begin{array}{cc} \tilde{A} & 0 \\ 0 & -\tilde{A} \end{array} \right].$$

Since the singular values of block-diagonal matrices are just the singular values of the blocks, we have $\Lambda_\epsilon(A) = \Lambda_\epsilon(\tilde{A}) \cup \Lambda_\epsilon(-\tilde{A})$, so we know $\Lambda_\epsilon^0(A) = \Lambda_\epsilon^0(\tilde{A}) \cup \Lambda_\epsilon^0(-\tilde{A})$. By considering a neighborhood of the point z_ϵ, this latter set cannot be convex.

In this example the difficulty is caused by the fact that the zero eigenvalue is derogatory. In what follows, we show good behavior of pseudospectra around nonderogatory eigenvalues, providing the parameter ϵ is sufficiently small.

3. Background results. We recall some results from [3]. A real-valued function on a real vector space is *real-analytic* at zero if in some neighborhood of zero it 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 C^∞ near the point in question.

The smallest singular value of the matrix Z is *simple* when the smallest eigenvalue of the Hermitian matrix Z^*Z is simple. Since the eigenvalues of matrices depend continuously on the matrix, the set of matrices Z with simple smallest singular values is open.

We consider the function $h : \mathbf{M}^n \times \mathbf{C} \to \mathbf{R}$ defined by

$$h(A, z) = (\sigma_{\min}(A - zI))^2.$$

For any $A \in \mathbf{M}^n$, we also define a function $h_A : \mathbf{C} \to \mathbf{R}$ by $h_A(z) = h(A, z)$. Treating $\mathbf{C}$ as a Euclidean space with inner product $\langle w, z \rangle = \operatorname{Re}(w^*z)$, we can interpret gradients $\nabla h_A(z)$ as elements of $\mathbf{C}$.

THEOREM 3.1 (analytic singular value). *If the smallest singular value of the matrix Z is simple, then the function $\sigma_{\min}^2$ is real-analytic at Z.*

An eigenvalue of A is *nonderogatory* if it has geometric multiplicity one. Among multiple eigenvalues, the nonderogatory ones are the most typical (from the perspective of the dimensions of the corresponding manifolds in $\mathbf{M}^n$ [1]). The matrix A is *nonderogatory* if all its eigenvalues are nonderogatory.

The following result is very well known.

PROPOSITION 3.2 (nonderogatory eigenvalues). *A matrix A has a nonderogatory eigenvalue λ if and only if zero is a simple singular value of $A - \lambda I$.*

The next result, an immediate consequence of [3, Thm. 7.4 and Cor. 7.8], shows that the resolvent norm is well-behaved near any nonderogatory eigenvalue of A. For a symmetric matrix X, we write $X \succ 0$ to mean X is positive-definite.

THEOREM 3.3 (growth near an eigenvalue). *Suppose λ is a nonderogatory eigenvalue of the matrix A. Then, for all $z \neq \lambda$ near λ, the function h_A is real-analytic with $\nabla h_A(z) \neq 0$ and $\nabla^2 h_A(z) \succ 0$.*

Related results appear in [6].

4. Convexity. In [3] we observe, as a consequence of Theorem 3.3 (growth near an eigenvalue), that if λ is a nonderogatory eigenvalue of a matrix A, then for small $\epsilon > 0$ the part of the pseudospectrum $\Lambda_\epsilon(A)$ near λ is strictly convex. (We call a closed set $S \subset \mathbf{C}$ *strictly convex* if the open line segment (u, v) lies in $\operatorname{int} S$ for any distinct points $u, v \in S$.) The first step in our development is to generalize this result to allow the matrix A to vary. We denote the closed unit disk in $\mathbf{C}$ by D and the closed unit ball in $\mathbf{M}^n$ by B.

We begin with a rather technical statement of our basic tool.

THEOREM 4.1 (small pseudospectra). *Consider a nonderogatory eigenvalue λ of a matrix $A_0 \in \mathbf{M}^n$. For any sufficiently small number $\mu > 0$, there exists a number $\bar{\epsilon} \in (0, \mu)$ (depending on μ) such that all numbers $\epsilon \in (0, \bar{\epsilon})$ have the following two properties.*

1. For all matrices $A \in \mathbf{M}^n$ in a neighborhood of A_0 (depending on μ and ϵ), the set

$$(4.1) \qquad \hat{\Lambda}_\epsilon(A) = \left\{ z \in \Lambda_\epsilon(A) : |z - \lambda| < \mu \right\}$$

is the component of the pseudospectrum $\Lambda_\epsilon(A)$ containing λ and contains no eigenvalues of A_0 except λ.

2. *There exists a number $\bar{\eta} \in (0, \mu)$ (depending on μ and ϵ) such that, given any number $\eta \in (0, \bar{\eta})$, all matrices A in a neighborhood of A_0 (depending on μ, ϵ, and η) satisfy, in addition to the above property,*

(i) *$\hat{\Lambda}_\epsilon(A)$ is compact, strictly convex, and contains $\lambda + \eta D$,*
and, for all points $z \in \lambda + \mu D$,
(ii) *the smallest singular value of $A - zI$ is simple, and*
(iii) *if $|z - \lambda| \geq \eta$, then $\nabla h_A(z) \neq 0$ and $\nabla^2 h_A(z) \succ 0$.*

Proof. Without loss of generality, $\lambda = 0$. By Theorem 3.3 (growth near an eigenvalue), there exists a number $\mu > 0$ such that

(4.2) $$0 < |z| \leq \mu \quad \Rightarrow \quad \nabla h_{A_0}(z) \neq 0 \text{ and } \nabla^2 h_{A_0}(z) \succ 0.$$

Hence the function h_{A_0} is strictly convex on the disk μD, with a strict local minimum value of zero at zero. In particular, we deduce

(4.3) $$\Lambda(A_0) \cap \mu D = \{0\}.$$

Consider the open set

$$\Omega = \Big\{(A, z) \in \mathbf{M}^n \times \mathbf{C} : \text{the smallest singular value of } A - zI \text{ is simple}\Big\}.$$

Theorem 3.1 (analytic singular value) implies that the function h is real-analytic throughout Ω, so the function $(A, z) \mapsto \nabla^2 h_A(z)$ is continuous on Ω. Clearly $(A_0, 0) \in \Omega$. Hence, by reducing μ if necessary, we can suppose there exists a number $\delta_1 > 0$ such that

$$\Big\{(A, z) \in \mathbf{M}^n \times \mu D : \|A - A_0\| < \delta_1\Big\} \subset \Omega.$$

Choose any number $\mu_1 \in (0, \mu)$. Then we claim

(4.4) $$\Lambda_\epsilon(A_0) \subset \mu_1 D \cup \mu D^c$$

for all small $\epsilon > 0$ (where D^c denotes the complement of D). If this were not the case, there would exist sequences of parameters $\epsilon_r \downarrow 0$ and points $z_r \in \Lambda_{\epsilon_r}(A_0)$ satisfying $\mu_1 < |z_r| \leq \mu$. By compactness, we can suppose z_r approaches a nonzero point $z \in \mu D$. However, since $\sigma_{\min}(A_0 - z_r I) \leq \epsilon_r$ for all r, we then deduce $\sigma_{\min}(A_0 - zI) \leq 0$, so $z \in \Lambda(A_0)$, contradicting (4.3).

Fix any $\epsilon > 0$ small enough to ensure inclusion (4.4), and choose any number $\mu_2 \in (\mu_1, \mu)$. We claim there exists a number $\delta_2 \in (0, \delta_1)$ such that

(4.5) $$\|A - A_0\| < \delta_2 \quad \Rightarrow \quad \hat{\Lambda}_\epsilon(A) \subset \mu_2 D.$$

Indeed, if this fails, there are sequences of matrices $A_r \to A_0$ and points $z_r \in \hat{\Lambda}_\epsilon(A_r)$ satisfying $\mu_2 < |z_r| < \mu$. By compactness, we can suppose z_r approaches a point $z \in \mu D$ satisfying $|z| \geq \mu_2 > \mu_1$. However, since $\sigma_{\min}(A_r - z_r I) \leq \epsilon$ for all r, we deduce $\sigma_{\min}(A_0 - zI) \leq \epsilon$, and hence $z \in \Lambda_\epsilon(A_0)$. But this contradicts inclusion (4.4).

The inclusion $\hat{\Lambda}_\epsilon(A) \subset \mu_2 D$ implies that the set $\hat{\Lambda}_\epsilon(A)$ is compact, being the intersection of the two compact sets $\Lambda_\epsilon(A)$ and $\mu_2 D$.

For our next step, observe that, by continuity, we know there exists a number $\eta \in (0, \mu)$ such that $\sigma_{\min}(A_0 - zI) < \epsilon$ for all points $z \in \eta D$. We now claim there exists a number $\delta_3 \in (0, \delta_2)$ such that

(4.6) $$\|A - A_0\| < \delta_3 \quad \Rightarrow \quad \eta D \subset \text{int } \hat{\Lambda}_\epsilon(A).$$

Suppose this property fails, so there are sequences of matrices $A_r \to A_0$ and points $z_r \in \eta D$ satisfying $z_r \notin \text{int } \hat{\Lambda}_\epsilon(A_r)$, and hence $\sigma_{\min}(A_r - z_r I) \geq \epsilon$. By compactness, we can suppose z_r approaches a point $z \in \eta D$, giving the contradiction $\sigma_{\min}(A_0 - zI) \geq \epsilon$.

We next claim there exists a number $\delta \in (0, \delta_3)$ such that, whenever $\|A - A_0\| \leq \delta$ and $\eta \leq |z| \leq \mu$, we have

$$(4.7) \qquad \nabla h_A(z) \neq 0 \quad \text{and} \quad \nabla^2 h_A(z) \succ 0.$$

If this fails, there are sequences of matrices $A_r \to A_0$ and points z_r satisfying $\eta \leq |z_r| \leq \mu$ and

$$\min\left\{ |\nabla h_{A_r}(z_r)|, \lambda_{\min}(\nabla^2 h_{A_r}(z_r)) \right\} \leq 0$$

for all r. By compactness, we can suppose z_r approaches a point $\hat{z}$ satisfying $\eta \leq |\hat{z}| \leq \mu$. By the continuity with respect to $(A, z) \in \Omega$ of the functions $\nabla h_A(z)$ and $\nabla^2 h_A(z)$, we deduce

$$\min\left\{ |\nabla h_{A_0}(\hat{z})|, \lambda_{\min}(\nabla^2 h_{A_0}(\hat{z})) \right\} \leq 0,$$

contradicting statement (4.2).

We next prove that the set $\hat{\Lambda}_\epsilon(A)$ is strictly convex. To this end, consider any matrix A satisfying $\|A - A_0\| < \delta$ and any two distinct points $u, v \in \hat{\Lambda}_\epsilon(A)$. We want to show the open line segment (u, v) lies in $\text{int } \hat{\Lambda}_\epsilon(A)$. By property (4.6), we know that

$$(4.8) \qquad \eta D \subset \text{int } \hat{\Lambda}_\epsilon(A).$$

We consider various cases.

(i) $|u|, |v| \leq \eta$. The result then follows by inclusion (4.8).

(ii) $(u, v) \cap \eta D = \emptyset$. In this case, we know $h_A(u) \leq \epsilon^2$ and $h_A(v) \leq \epsilon^2$, and the function h_A is strictly convex on the line segment $[u, v]$, by property (4.7), so the result follows.

(iii) $|u| \leq \eta$ and $|v| > \eta$. Then consider the unique number $\gamma \in [0, 1]$ such that the point $w = \gamma u + (1 - \gamma)v$ satisfies $|w| = \eta$. Then $[u, w] \subset \text{int } \hat{\Lambda}_\epsilon(A)$ by inclusion (4.8), while $(w, v) \subset \text{int } \hat{\Lambda}_\epsilon(A)$ by case (ii).

(iv) $|u| > \eta$ and $|v| \leq \eta$. By swapping u and v, we obtain case (iii).

(v) $|u|, |v| > \eta$ and $(u, v) \cap \eta D \neq \emptyset$. Consider the two (possibly equal) solutions $\gamma_1 \geq \gamma_2$ in $[0, 1]$ to the quadratic equation $|\gamma u + (1 - \gamma)v|^2 = \eta^2$. For each $j = 1, 2$, set $w_j = \gamma_j u + (1 - \gamma_j)v$. Then $[w_1, w_2] \subset \text{int } \hat{\Lambda}_\epsilon(A)$ by inclusion (4.8), while both intervals (u, w_1) and (w_2, v) lie in $\text{int } \hat{\Lambda}_\epsilon(A)$ by case (ii).

This completes the proof of strict convexity.

To see that the set $\hat{\Lambda}_\epsilon(A)$ must be the component of the pseudospectrum $\Lambda_\epsilon(A)$ containing λ, note that the function $A \mapsto \sigma_{\min}(A - \lambda I)$ is continuous on $\mathbf{M}^n$, and $\sigma_{\min}(A_0 - \lambda I) = 0$, so $\lambda \in \hat{\Lambda}_\epsilon(A)$ for all A near A_0. Since the $\hat{\Lambda}_\epsilon(A)$ is a connected subset of $\Lambda_\epsilon(A)$, being convex, the result follows. $\square$

COROLLARY 4.2 (strict convexity). *Consider a nonderogatory eigenvalue λ of a matrix $A_0 \in \mathbf{M}^n$. Given any sufficiently small $\epsilon > 0$, the component of the pseudospectrum $\Lambda_\epsilon(A)$ containing λ is strictly convex for all matrices A sufficiently close to A_0.*

5. Sensitivity. We are now ready to study the dependence of a fixed component of the pseudospectrum $\Lambda_\epsilon(A)$ on the matrix A.

LEMMA 5.1 (gradient continuity). *Suppose all the assumptions of Theorem* 4.1 *(small pseudospectra) hold. For nonzero complex w, consider the function $\alpha_w : \mathbf{M}^n \to \mathbf{R}$ defined by*

$$\alpha_w(A) = \sup\{\operatorname{Re}(w^* z) : z \in \hat{\Lambda}_\epsilon(A)\}. \tag{5.1}$$

Then the function $(A, w) \mapsto \alpha_w(A)$ is C^∞ on the set

$$\Big\{(A, w) \in \mathbf{M}^n \times \mathbf{C} : \|A - A_0\| \leq \delta, \ w \neq 0\Big\}.$$

Proof. The supremum (5.1) is attained at a unique point $z(A, w) \in \hat{\Lambda}_\epsilon(A)$, since the set $\hat{\Lambda}_\epsilon(A)$ is compact and strictly convex. We can also write the supremum as a smooth optimization problem,

$$\alpha_w(A) = \sup \Big\{\operatorname{Re}(w^* z) : h_A(z) \leq \epsilon^2, \ |z - \lambda| < \mu\Big\}.$$

By continuity, the optimal solution $z(A, w)$ must satisfy $h_A(z(A, w)) = \epsilon^2$. The function h_A is real-analytic (so, in particular, C^∞), and satisfies the condition

$$\nabla h_A(z(A, w)) \neq 0 \ \text{ and } \ \nabla^2 h_A(z(A, w)) \succ 0. \tag{5.2}$$

We now apply a standard sensitivity argument to show that the dependence of the optimal solution $z(A, w)$ on the parameters (A, w) is also C^∞. We argue as follows. Since $\nabla h_A(z(A, w)) \neq 0$, there exists a Lagrange multiplier $\gamma(A, w) \in \mathbf{R}$ corresponding to the optimal solution. Thus $z = z(A, w)$ and $\gamma = \gamma(A, w)$ solve the system

$$w + \gamma \nabla h_A(z) = 0,$$
$$h_A(z) = \epsilon^2.$$

But it is easy to check that condition (5.2) implies that the Jacobian for the left-hand side is surjective at $(z(A, w), \gamma(A, w))$. Hence the implicit function theorem implies that the mapping $(A, w) \mapsto z(A, w)$ is C^∞. The result follows. $\square$

We can now prove our main result.

THEOREM 5.2 (component Lipschitz behavior). *Consider a nonderogatory eigenvalue λ of a matrix $A_0 \in \mathbf{M}^n$. For any sufficiently small number $\mu > 0$, there exists a number $\bar{\epsilon} \in (0, \mu)$ (depending on μ) such that for all numbers $\epsilon \in (0, \bar{\epsilon})$, and all matrices $A \in \mathbf{M}^n$ in a neighborhood of A_0 (depending on μ and ϵ), the set*

$$\hat{\Lambda}_\epsilon(A) \ = \ \Big\{z \in \Lambda_\epsilon(A) : |z - \lambda| < \mu\Big\}$$

has the following properties:

(i) *$\hat{\Lambda}_\epsilon(A)$ is the component of the pseudospectrum $\Lambda_\epsilon(A)$ containing λ.*
(ii) *$\hat{\Lambda}_\epsilon(A)$ contains no eigenvalues of A_0 except λ.*
(iii) *$\hat{\Lambda}_\epsilon(A)$ is compact and strictly convex.*
(iv) *The set-valued mapping $\hat{\Lambda}_\epsilon$ is Lipschitz on a neighborhood of A_0 (with respect to the Hausdorff distance).*

Proof. We apply Theorem 4.1 (small pseudospectra) and Corollary 4.2 (strict convexity). Using Lemma 5.1 (gradient continuity), we can define a number

$$L = \max\left\{ \|\nabla\alpha_w(A)\| : A \in \Gamma, \ |w| = 1 \right\},$$

where the set Γ is the neighborhood of the matrix A_0 referred to in Theorem 4.1. Consider any two matrices $A_1, A_2 \in \Gamma$. According to [7, Lemma 2], the Hausdorff distance between the sets $\hat{\Lambda}_\epsilon(A_1)$ and $\hat{\Lambda}_\epsilon(A_2)$ is given by

$$\max_{|w|=1} \left| \alpha_w(A_1) - \alpha_w(A_2) \right|,$$

and, by the definition of L, this quantity cannot exceed $L\|A_1 - A_2\|$. $\quad\square$

In particular, we obtain the following variational property of the entire pseudopectrum.

COROLLARY 5.3 (pseudospectral Lipschitz behavior). *If the matrix $A_0 \in \mathbf{M}^n$ is nonderogatory, then for any small $\epsilon > 0$, the dependence of the pseudospectrum $\Lambda_\epsilon(A)$ on the matrix $A \in \mathbf{M}^n$ is Lipschitz near A_0 (with respect to the Hausdorff distance).*

Proof. Denote the distinct eigenvalues of A_0 by $\lambda_1, \lambda_2, \ldots, \lambda_m$, and denote the separation of the eigenvalues by $\nu = \min_{j \neq k} |\lambda_j - \lambda_k|$. Now apply the preceding result successively at each eigenvalue λ_j to obtain a number $\mu < \nu/3$ such that any small $\epsilon > 0$ has the following property: for all matrices A near A_0 and each index $j = 1, 2, \ldots, m$, the component of the pseudospectrum $\Lambda_\epsilon(A)$ containing λ_j is

$$\Lambda_\epsilon^j(A) = \{z \in \Lambda_\epsilon(A) : |z - \lambda_j| < \mu\},$$

and the set-valued mapping Λ_ϵ^j is Lipschitz around A_0.

Now consider any matrices $A_1, A_2 \in \mathbf{M}^n$ near A_0. For any fixed index j, we have

$$(5.3) \qquad z \in \Lambda_\epsilon^j(A_1) \quad \Rightarrow \quad d(z, \Lambda_\epsilon(A_2)) = d(z, \Lambda_\epsilon^j(A_2)).$$

To see this, notice that $d(z, \Lambda_\epsilon^j(A_2)) < \mu$ because $\lambda_j \in \Lambda_\epsilon^j(A_2)$. On the other hand, for indices $k \neq j$, we know

$$|z - \lambda_j| < \mu, \quad |\lambda_j - \lambda_k| > 3\mu, \quad \Lambda_\epsilon^k(A_2) \subset \lambda_k + \mu D,$$

so $d(z, \Lambda_\epsilon^k(A_2)) > \mu$. Since

$$d(z, \Lambda_\epsilon(A_2)) = \min_k d(z, \Lambda_\epsilon^k(A_2)),$$

our claim (5.3) now follows.

As a consequence of the implication (5.3), we obtain

$$\sup_{z \in \Lambda_\epsilon(A_1)} d(z, \Lambda_\epsilon(A_2)) = \max_j \sup_{z \in \Lambda_\epsilon^j(A_1)} d(z, \Lambda_\epsilon(A_2)) = \max_j \sup_{z \in \Lambda_\epsilon^j(A_1)} d(z, \Lambda_\epsilon^j(A_2)),$$

and similarly,

$$\sup_{z \in \Lambda_\epsilon(A_2)} d(z, \Lambda_\epsilon(A_1)) = \max_k \sup_{z \in \Lambda_\epsilon^k(A_2)} d(z, \Lambda_\epsilon^k(A_1)).$$

Hence the Hausdorff distance between the pseudospectra $\Lambda_\epsilon(A_1)$ and $\Lambda_\epsilon(A_2)$ is given by

$$
\begin{aligned}
H\Big(\Lambda_\epsilon(A_1), \Lambda_\epsilon(A_2)\Big) &= \max\left\{ \sup_{z\in\Lambda_\epsilon(A_1)} d(z, \Lambda_\epsilon(A_2)), \sup_{z\in\Lambda_\epsilon(A_2)} d(z, \Lambda_\epsilon(A_1)) \right\} \\
&= \max\left\{ \max_j \sup_{z\in\Lambda_\epsilon^j(A_1)} d(z, \Lambda_\epsilon^j(A_2)), \max_k \sup_{z\in\Lambda_\epsilon^k(A_2)} d(z, \Lambda_\epsilon^k(A_1)) \right\} \\
&= \max_r \max\left\{ \sup_{z\in\Lambda_\epsilon^r(A_1)} d(z, \Lambda_\epsilon^r(A_2)), \sup_{z\in\Lambda_\epsilon^r(A_2)} d(z, \Lambda_\epsilon^r(A_1)) \right\} \\
&= \max_r H\Big(\Lambda_\epsilon^r(A_1), \Lambda_\epsilon^r(A_2)\Big).
\end{aligned}
$$

The result now follows.   □

Acknowledgments. The authors thank Mark Embree and Nick Trefethen for a number of insights that improved the overall presentation.

Note added in proof. A generalization of Corollary 5.3 to the derogatory case appears in a recent preprint, "Variational Analysis of Pseudospectra" by A. S. Lewis and C. H. J. Pang.

REFERENCES

[1] V. I. ARNOLD, *On matrices depending on parameters*, Uspehi Mat. Nauk, 26 (1971), pp. 29–43.
[2] R. BENEDETTI AND J.-J. RISLER, *Real Algebraic and Semi-Algebraic Sets*, Hermann, Paris, 1990.
[3] J. V. BURKE, A. S. LEWIS, AND M. L. OVERTON, *Optimization and pseudospectra, with applications to robust stability*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 80–104. Corrigendum: www.cs.nyu.edu/cs/faculty/overton/papers/pseudo_corrigendum.html.
[4] J. V. BURKE, A. S. LEWIS, AND M. L. OVERTON, *Spectral conditioning and pseudospectral growth*, Numer. Math., to appear.
[5] M. EMBREE AND L. N. TREFETHEN, *Pseudospectra Gateway*, web.comlab.ox.ac.uk/pseudospectra.
[6] M. KAROW, *Eigenvalue condition numbers and a formula of Burke, Lewis and Overton*, Electron. J. Linear Algebra, 15 (2006), pp. 143–153.
[7] G. SALINETTI AND R. J.-B. WETS, *On the convergence of sequences of convex sets in finite dimensions*, SIAM Rev., 21 (1979), pp. 18–33.
[8] L. N. TREFETHEN, *private communication*, 2005.
[9] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, NJ, 2005.