Electronic Transactions on Numerical Analysis.
Volume 38, pp. 168-183, 2011.

ISSN 1068-9613.

ETNA
Kent State University
http://etna.math.kent.edu

PSEUDOSPECTRAL MAPPING THEOREM II*

S. H. LUI†

Abstract. The pseudospectrum has become an important quantity for analyzing stability of non-normal systems. This paper is a continuation of an earlier paper of this author where a mapping theorem for pseudospectra was given, generalizing the spectral mapping theorem for eigenvalues. The main contribution of this paper consists of asymptotic expansions of quantities which determine the sizes of components of pseudospectral sets. As an application of this theory, we solve the eigenvalue perturbation problem for an analytic function of a matrix. Some numerical examples illustrate the theory.

Key words. Eigenvalues, pseudospectra, spectral mapping theorem, condition number, eigenvalue perturbation of function of matrices

AMS subject classifications. 15A18, 15A60, 65F15

1. Introduction.

The properties of a normal matrix can be accurately predicted by its spectrum. Here, normality refers to the matrix having a complete set of orthogonal eigenvectors. The spectrum of a non-normal matrix, however, may not be very informative. Thanks largely to the work of Trefethen and his co-workers, the pseudospectrum has emerged as an appropriate indicator for the stability of non-normal systems. It has been applied to problems in hydrodynamic instability, turbulence, magnetohydrodynamics, control theory, iterative solution of linear equations, numerical solution of differential equations, quantum mechanics, random matrices, etc. See [6] for an authoritative survey and references and [4] for an exposition of classical eigenvalue perturbation theory.

For a square matrix A and a non-negative number ϵ, the ϵ-pseudospectrum of A is defined as the following closed set in the complex plane:

$$\Lambda_\epsilon(A) \equiv \bigcup_{\|E\|\le\epsilon} \Lambda(A+E).$$

Here $\Lambda()$ denotes the spectrum of a matrix and $\|\cdot\|$ is the matrix 2-norm. (This definition is slightly different from that given in [6] where the inequality is replaced by strict inequality.) An equivalent definition is

$$\Lambda_\epsilon(A) \equiv \{z \in \mathbb{C},\ \|(zI-A)^{-1}\| \ge \epsilon^{-1}\}$$

where the norm is taken to be infinite if $z \in \Lambda(A)$. When A is a normal matrix, its ϵ-pseudospectrum is the union of closed disks of radius ϵ with centers at the eigenvalues. For a non-normal matrix, its ϵ-pseudospectrum can be much bigger than this union.

The spectral mapping theorem is a fundamental result in functional analysis of great importance. Given a matrix A and a function f which is analytic on an open set containing $\Lambda(A)$, the theorem asserts that

$$f(\Lambda(A)) = \Lambda(f(A)).$$

In [5], we discussed a mapping theorem for ϵ-pseudospectrum which generalizes the spectral mapping theorem in the sense that when $\epsilon = 0$, the pseudospectral mapping theorem becomes the spectral mapping theorem.

* Received July 26, 2010. Accepted for publication May 5, 2011. Published online June 14, 2011. Recommended by F. Dopico. This work was in part supported by a grant from NSERC.
†Department of Mathematics, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2 (luish@cc.umanitoba.ca).

ETNA
Kent State University
http://etna.math.kent.edu

THEOREM 1.1 (pseudospectral mapping theorem). *Let A be a matrix and f be an analytic function defined on an open set containing $\Lambda(A)$. For each ϵ, $s \geq 0$ sufficiently small, define*

$$\phi(\epsilon) = \sup_{\zeta \in \Lambda_\epsilon(A)} \inf\{r \geq 0,\ f(\zeta) \in \Lambda_r(f(A))\}$$

and

$$\psi(s) = \sup_{z \in \Lambda_s(f(A))} \inf\{r \geq 0,\ z \in f(\Lambda_r(A))\}.$$

Then

$$f(\Lambda_\epsilon(A)) \subset \Lambda_{\phi(\epsilon)}(f(A)) \subset f(\Lambda_{\psi(\phi(\epsilon))}(A)).$$

Let A be a matrix with distinct eigenvalues $\{\lambda_j,\ j = 1, \cdots, k\}$ each having some positive algebraic multiplicity. When ϵ is small, $\Lambda_\epsilon(A)$ consists of k disjoint components each containing an eigenvalue. These components are approximately disks ([3]). In the pseudospectral mapping theorem, the sizes of pseudospectra are characterized by one pair of functions ϕ and ψ. Our first order of business is to characterize each component by functions ϕ_j and ψ_j, offering a sharper bound than the one in the pseudospectral mapping theorem. While the functions ϕ_j and ψ_j are continuous and monotonically increasing, it appears to be difficult to derive other properties. The main purpose of this paper is to obtain the first term in the asymptotic expansions of ϕ_j and ψ_j.

In Section 2, we derive the exact expressions for ϕ_j and ψ_j mentioned in the previous paragraph. In Section 3, we determine the size of each component of the pseudospectrum of $f(A)$. This is followed by a derivation of the asymptotic expansions. In Section 5, we apply these results to obtain sharp estimates for how the eigenvalues of $f(A)$ perturb when there is a perturbation in A. In fact, we estimate the condition number of the eigenvalue $f(\lambda_j)$ of $f(A)$ when A is subject to a perturbation. Some numerical experiments in the final section illustrate the theory.

2. A component-wise pseudospectral mapping theorem.

The following is a sharper version of the pseudospectral mapping theorem for complex analytic functions discussed in [5]. The proof is the same as that in [5] for the original theorem and is included here for completeness.

As already mentioned in the introduction, when ϵ is small, $\Lambda_\epsilon(A)$ is a disjoint union of sets each containing exactly one eigenvalue. Denote the component containing the distinct eigenvalue λ_j by $\Lambda_\epsilon(A, \lambda_j)$. *Throughout this paper, we shall be assuming that the parameter ϵ is sufficiently small so that the components of pseudospectral sets are pairwise disjoint.* The value of ϵ may need to be restricted further. This point will be elaborated upon later. The same assumption applies to the parameter s used in the context of pseudospectral sets for $f(A)$. In case $f(\lambda_j) = f(\lambda_k)$ for some $\lambda_k \neq \lambda_j$, we identify the two components $\Lambda_s(f(A), f(\lambda_j))$ and $\Lambda_s(f(A), f(\lambda_k))$ for all $s \geq 0$.

THEOREM 2.1. *Let A be a matrix with eigenvalues $\{\lambda_j\}$ and f be an analytic function defined on an open set containing $\Lambda(A)$. For each j and each ϵ, $s \geq 0$ sufficiently small, define*

$$\phi_j(\epsilon) = \sup_{\zeta \in \Lambda_\epsilon(A, \lambda_j)} \inf\{r \geq 0,\ f(\zeta) \in \Lambda_r(f(A), f(\lambda_j))\}$$

ETNA
Kent State University
http://etna.math.kent.edu

and

$$\psi_j(s) = \sup_{z\in\Lambda_s(f(A),f(\lambda_j))} \inf\{r \geq 0,\ z \in f(\Lambda_r(A,\lambda_j))\}.$$

Then

$$f(\Lambda_\epsilon(A,\lambda_j)) \subset \Lambda_{\phi_j(\epsilon)}(f(A),f(\lambda_j)) \subset f(\Lambda_{\psi_j(\phi_j(\epsilon))}(A,\lambda_j)).$$

Proof. Fix some j. We first show that ϕ_j is well defined. Let $\zeta \in \Lambda_\epsilon(A,\lambda_j)$. Then $\zeta \in \Lambda(A+E)$ for some matrix E such that $\|E\| \leq \epsilon$. By the spectral mapping theorem,

$$f(\zeta) \in \Lambda(f(A+E)) = \Lambda(f(A)+F)$$

where $F = f(A+E) - f(A)$. Thus $f(\zeta) \in \Lambda_{\|F\|}(f(A),f(\lambda_j))$ which implies that the infimum in the definition of ϕ_j is taken over a non-empty set and thus ϕ_j is well defined. The first set inclusion now follows directly from the definition of ϕ_j.

Next, we show that ψ_j is well defined assuming that f is not a constant. (If f is a constant, then $\phi_j \equiv 0$ and $\psi_j(0) = 0$ and the theorem is trivially true.) Let $z \in \Lambda_s(f(A),f(\lambda_j))$ for some small positive s. By the Open Mapping Theorem of complex analysis, there are some $r > 0$ and $\zeta \in B_r(\lambda_j)$, the open disk of radius r and center λ_j, so that $z = f(\zeta)$. (Note that s must be so small that the Open Mapping Theorem is applicable to f as a mapping from $B_r(\lambda_j)$ to some open set containing $\Lambda_s(f(A),f(\lambda_j))$.) Since $B_r(\lambda_j) \subset \Lambda_r(A,\lambda_j)$, we have $z \in f(\Lambda_r(A,\lambda_j))$. Thus, the infimum in the definition of ψ_j is taken over a non-empty set and so ψ_j is well defined. The second set inclusion now follows directly from the definition of $\psi_j(s)$ with $s = \phi_j(\epsilon)$. (Note that the value of ϵ may need to be reduced so that $s = \phi_j(\epsilon)$ is small.) $\square$

An equivalent conclusion to the above theorem is that, for small s,

$$\Lambda_s(f(A),f(\lambda_j)) \subset f(\Lambda_{\psi_j(s)}(A,\lambda_j)) \subset \Lambda_{\phi_j(\psi_j(s))}(f(A),f(\lambda_j)). \tag{2.1}$$

Note that by the definitions of ϕ_j and ψ_j, the set inclusions are sharp in the sense that the functions cannot be replaced by smaller functions.

3. The size of the pseudospectral component of $f(A)$. In this section, we estimate the size of each component of the pseudospectrum of $f(A)$ where f is analytic. An eigenvalue is semi-simple if its algebraic multiplicity coincides with its geometric multiplicity. The $m\times m$ identity matrix is denoted by I_m. For any set S, the boundary of the set is denoted by ∂S. The following is a translation of a classical result (see, p. 69 in [7] and [3, Theorem 3.1]) to the language of pseudospectra.

THEOREM 3.1. *Suppose λ is a semi-simple eigenvalue of A of multiplicity $m \geq 1$. Let $A = QJQ^{-1}$ where*

$$J = \begin{bmatrix} \lambda I_m & \\ & J_2 \end{bmatrix}$$

is a Jordan form of A with $\lambda \notin \Lambda(J_2)$. Let $\epsilon > 0$. For any $z \in \partial\Lambda_\epsilon(A,\lambda)$,

$$|z-\lambda| = \epsilon\|P\| + O(\epsilon^2)$$

where P is the projection onto the eigenspace ker $(A-\lambda I)$ along the range space of $A-\lambda I$:

$$P = Q\begin{bmatrix} I_m & \\ & 0 \end{bmatrix} Q^{-1}. \tag{3.1}$$

ETNA
Kent State University
http://etna.math.kent.edu

Proof. Let $z \in \partial\Lambda_\epsilon(A, \lambda)$. Observe that

$$(zI - A)^{-1} = Q \begin{bmatrix} (z-\lambda)^{-1} I_m & \\ & (zI - J_2)^{-1} \end{bmatrix} Q^{-1}.$$

Now,

$$\begin{aligned} \frac{1}{\epsilon} = \|(zI - A)^{-1}\| &= \left\| Q \begin{bmatrix} (z-\lambda)^{-1} I_m & \\ & 0 \end{bmatrix} Q^{-1} + Q \begin{bmatrix} 0 & \\ & (zI - J_2)^{-1} \end{bmatrix} Q^{-1} \right\| \\ &= \frac{\|P\|}{|z-\lambda|} + O(1). \end{aligned}$$

This implies that

$$|z - \lambda| = \frac{\epsilon\|P\|}{1 + O(\epsilon)} = \epsilon\,\|P\| + O(\epsilon^2). \qquad \square$$

We remark that in case λ is a simple eigenvalue, then it is well known that $P = \frac{xy^*}{y^*x}$ where x and y are right and left, respectively, eigenvectors corresponding to λ.

COROLLARY 3.2. *Suppose A is diagonalizable: $A = QDQ^{-1}$ for some diagonal D. Assume f is analytic on some open set containing $\Lambda(A)$. Let λ be any eigenvalue of A and $\tilde{m}$ be the multiplicity of $f(\lambda)$ as an eigenvalue of $f(A)$. Define*

$$\tilde{P} = Q \begin{bmatrix} I_{\tilde{m}} & \\ & 0 \end{bmatrix} Q^{-1} \tag{3.2}$$

assuming all eigenvalues μ so that $f(\mu) = f(\lambda)$ are placed in the first $\tilde{m}$ diagonal entries of D. Let $s > 0$. Then for any $\zeta \in \partial\Lambda_s(f(A), f(\lambda))$,

$$|\zeta - f(\lambda)| = s\,\|\tilde{P}\| + O(s^2). \tag{3.3}$$

Proof. Note that

$$f(D) = \begin{bmatrix} f(\lambda) I_{\tilde{m}} & \\ & f(D_2) \end{bmatrix}$$

where D_2 is diagonal so that $f(\mu)$ is distinct from $f(\lambda)$ for any diagonal entry μ of D_2. The result now follows from a direct application of Theorem 3.1. $\square$

In Corollary 3.2, suppose λ is an eigenvalue of A of multiplicity m. If A has an eigenvalue μ distinct from λ so that $f(\mu) = f(\lambda)$, then $\tilde{m} > m$. Otherwise, $\tilde{m} = m$.

The index of an eigenvalue is the size of the largest Jordan block of that eigenvalue. The following theorem is very similar to results in the literature ([1, Theorem 7.4], (2.8) in [2] and [3, Theorem 3.1]).

THEOREM 3.3. *Suppose λ is an eigenvalue of A of index $m > 1$ and there is exactly one Jordan block associated with λ of size m. Let $A = QJQ^{-1}$ where*

$$J = \left[\begin{array}{ccccc|c} \lambda & 1 & & & & \\ & \ddots & \ddots & & & \\ & & \lambda & 1 & & \\ & & & \lambda & & \\ \hline & & & & & J_2 \end{array}\right] \tag{3.4}$$

ETNA
Kent State University
http://etna.math.kent.edu

is a Jordan form of A with the first block $m \times m$. Let $\epsilon > 0$. For any $z \in \partial\Lambda_\epsilon(A, \lambda)$,

$$|z - \lambda| = \epsilon^{1/m} \|N^{m-1}\|^{1/m} + O(\epsilon^{2/m})$$

where N is the nilpotent matrix associated with λ in the above Jordan decomposition of A:

$$N = Q \left[\begin{array}{cccc|c} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & 0 & 1 & \\ & & & 0 & \\ \hline & & & & 0 \end{array}\right] Q^{-1}. \tag{3.5}$$

Proof. Let $z \in \partial\Lambda_\epsilon(A, \lambda)$. Observe that

$$(zI - A)^{-1} = Q \left[\begin{array}{cccc|c} (z-\lambda)^{-1} & (z-\lambda)^{-2} & \cdots & (z-\lambda)^{-m} & \\ & \ddots & \ddots & \vdots & \\ & & (z-\lambda)^{-1} & (z-\lambda)^{-2} & \\ & & & (z-\lambda)^{-1} & \\ \hline & & & & (zI - J_2)^{-1} \end{array}\right] Q^{-1}.$$

Since the leading order term in the above matrix is $(z - \lambda)^{-m}$,

$$\begin{aligned} \frac{1}{\epsilon} &= \|(zI - A)^{-1}\| \\ &= |z - \lambda|^{-m} \left\| Q \left[\begin{array}{cccc|c} 0 & \cdots & 0 & 1 & \\ & 0 & \cdots & 0 & \\ & & \ddots & \vdots & \\ & & & 0 & \\ \hline & & & & 0 \end{array}\right] Q^{-1} \right\| + O(|z - \lambda|^{1-m}) \\ &= |z - \lambda|^{-m} \|N^{m-1}\| + O(|z - \lambda|^{1-m}). \end{aligned}$$

This implies that

$$|z - \lambda|^m = \epsilon \, \|N^{m-1}\| + O(\epsilon \, |z - \lambda|)$$

and the result now follows. ∎

In this theorem, we assume for ease of exposition that there is only one Jordan block of size m for the eigenvalue λ. The result also holds if there are $k \geq 1$ such Jordan blocks. In this case, the first diagonal block in (3.5) must be replicated k times.

COROLLARY 3.4. *Assume the hypotheses of the above theorem. Let f be analytic on some open set containing $\Lambda(A)$ so that $f'(\lambda) \neq 0$. Suppose $f(\lambda) \neq f(\mu)$ for every eigenvalue μ of A distinct from λ. Let $s > 0$. For any $\zeta \in \partial\Lambda_s(f(A), f(\lambda))$,*

$$|\zeta - f(\lambda)| = s^{1/m} \, |f'(\lambda)|^{1-\frac{1}{m}} \, \|N^{m-1}\|^{1/m} + O(s^{2/m}).$$

Proof. Since $\zeta \in \partial\Lambda_s(f(A), f(\lambda))$,

$$\frac{1}{s} = \|(\zeta I - f(A))^{-1}\| = \|Q(\zeta I - f(J))^{-1} Q^{-1}\|.$$

ETNA
Kent State University
http://etna.math.kent.edu

Let J_1 be the first diagonal block of (3.4) and N_1 be the $m \times m$ nilpotent matrix which is zero everywhere except for ones along the first superdiagonal. Recall that

$$f(J_1) = f(\lambda)I_m + f'(\lambda)N_1 + \cdots + \frac{f^{(m-1)}(\lambda)}{(m-1)!}N_1^{m-1} = \begin{bmatrix} f(\lambda) & f'(\lambda) & \cdots & \frac{f^{(m-1)}(\lambda)}{(m-1)!} \\ & \ddots & \ddots & \vdots \\ & & f(\lambda) & f'(\lambda) \\ & & & f(\lambda) \end{bmatrix}.$$

The matrix $\zeta I_m - f(J_1)$ can be explicitly inverted and we find that the dominant term which appears in the top right corner is

$$f'(\lambda)^{m-1}\delta^{-m} + O(|\delta|^{1-m}) \tag{3.6}$$

where $\delta = \zeta - f(\lambda)$. Hence,

$$\frac{1}{s} = |\delta|^{-m}|f'(\lambda)|^{m-1}\|N^{m-1}\| + O(|\delta|^{1-m})$$

which leads to

$$|\delta|^m = s\,|f'(\lambda)|^{m-1}\,\|N^{m-1}\| + O(s\,|\delta|)$$

from which the desired result follows. ☐

We next indicate briefly what happens in case some of the hypotheses in the above fail. For instance, assume $f(\lambda) = f(\mu)$ for some eigenvalue μ with largest index $\tilde{m} > m$. Suppose $f'(\mu) \neq 0$. Then the dominant behaviour comes from the Jordan block corresponding to μ of dimension $\tilde{m}$. In this case, we obtain

$$|\zeta - f(\lambda)| = s^{1/\tilde{m}}\,|f'(\mu)|^{1-\frac{1}{\tilde{m}}}\|\tilde{N}^{\tilde{m}-1}\|^{1/\tilde{m}} + O(s^{2/\tilde{m}}), \qquad \zeta \in \partial\Lambda_s(f(A), f(\lambda)),$$

where $\tilde{N}$ is the nilpotent matrix associated with the Jordan block of μ of size $\tilde{m}$.

Next, assume that the hypotheses of Corollary 3.4 holds, except that $f'(\lambda) = 0$ and $f''(\lambda) \neq 0$. First assume that the index of λ is odd: $m = 2k+1$. It can be checked that the dominant term of $(\zeta I_m - f(J_1))^{-1}$ again occurs in the top right corner and is $2^{-k}f''(\lambda)^k\delta^{-k-1} + O(|\delta|^{-k})$ where $\delta = \zeta - f(\lambda)$. This leads to

$$|\zeta - f(\lambda)| = s^{1/(k+1)}\left(\frac{|f''(\lambda)|}{2}\right)^{k/(k+1)}\|N^{m-1}\|^{1/(k+1)} + O(s^{2/(k+1)}),$$
$$\zeta \in \partial\Lambda_s(f(A), f(\lambda)).$$

If the index of λ is even: $m = 2k$, then the dominant term of $(\zeta I_m - f(J_1))^{-1}$ is $O(|\delta|^{-k})$ and it occurs at the $(1, m-1)$, $(2, m)$ and $(1, m)$ entries of the matrix if $m \geq 4$. If $m = 2$, then the dominant term occurs at the $(1, 2)$ entry.

4. Asymptotic expansions. In this section, we give asymptotic expansions for the functions ϕ_j and ψ_j in Theorem 2.1. We first discuss the case of a diagonalizable matrix.

THEOREM 4.1. *Let λ_1 be an eigenvalue of multiplicity $m \geq 1$ of a diagonalizable matrix $A = QDQ^{-1}$ where D is diagonal:*

$$D = \begin{bmatrix} \lambda_1 I_m & \\ & D_2 \end{bmatrix}$$

ETNA
Kent State University
http://etna.math.kent.edu

and D_2 is diagonal with $\lambda_1 \notin \Lambda(D_2)$. Let f be a function which is analytic in some open set containing $\Lambda(A)$. For small $\epsilon > 0$,

$$\phi_1(\epsilon) = \begin{cases} |f'(\lambda_1)| \frac{\|P\|}{\|\tilde{P}\|}\, \epsilon + O(\epsilon^2), & f'(\lambda_1) \neq 0; \\ \frac{|f''(\lambda_1)|}{2} \frac{\|P\|^2}{\|\tilde{P}\|}\, \epsilon^2 + O(\epsilon^3), & f'(\lambda_1) = 0,\ f''(\lambda_1) \neq 0 \end{cases}$$

where P and $\tilde{P}$ are as defined in (3.1) and (3.2) with $\tilde{m}$ the multiplicity of $f(\lambda_1)$ as an eigenvalue of $f(A)$. For small $s > 0$,

$$\psi_1(s) = \begin{cases} \frac{s}{|f'(\lambda_1)|} \frac{\|\tilde{P}\|}{\|P\|} + O(s^2), & f'(\lambda_1) \neq 0; \\ \frac{\sqrt{2}\,\|\tilde{P}\|^{1/2}\, s^{1/2}}{|f''(\lambda_1)|^{1/2}\|P\|} + O(s), & f'(\lambda_1) = 0,\ f''(\lambda_1) \neq 0. \end{cases}$$

Proof. By definition,

$$\begin{aligned} \phi_1(\epsilon) &= \sup_{z \in \Lambda_\epsilon(A,\lambda_1)} \inf\{r > 0,\ f(z) \in \Lambda_r(f(A), f(\lambda_1))\} \\ &= \sup_{z \in \Lambda_\epsilon(A,\lambda_1)} \inf\{r > 0,\ \|(f(z)I - f(A))^{-1}\| \geq r^{-1}\} \\ &= \sup_{z \in \Lambda_\epsilon(A,\lambda_1)} \|(f(z)I - f(A))^{-1}\|^{-1} \\ &= \sup_{z \in \Lambda_\epsilon(A,\lambda_1)} \|\, [Q(f(z)I - f(D))Q^{-1}]^{-1}\|^{-1} \\ &= \sup_{z \in \Lambda_\epsilon(A,\lambda_1)} \|Q(f(z)I - f(D))^{-1}Q^{-1}\|^{-1}. \end{aligned}$$

Define $\delta = f(z) - f(\lambda_1)$ which has a small magnitude when $z \in \Lambda_\epsilon(A, \lambda_1)$. Note that

$$f(z)I - f(D) = \begin{bmatrix} \delta I_{\tilde{m}} & \\ & f(z)I - f(D_3) \end{bmatrix},$$

where D_3 is diagonal so that $f(\mu) \neq f(\lambda_1)$ for every diagonal entry μ of D_3. Hence,

$$\|\, Q(f(z)I - f(D))^{-1}Q^{-1}\| = \frac{1}{|\delta|}\|\tilde{P}\| + O(1).$$

If $f'(\lambda_1) \neq 0$, then $\delta = f'(\lambda_1)(z - \lambda_1) + O(|z - \lambda_1|^2) = f'(\lambda_1)\,\epsilon\,\|P\| + O(\epsilon^2)$ by Theorem 3.1. Hence,

$$\phi_1(\epsilon) = \frac{|f'(\lambda_1)|\,\|P\|}{\|\tilde{P}\|}\epsilon + O(\epsilon^2).$$

Now assume that $f'(\lambda_1) = 0$ and $f''(\lambda_1) \neq 0$. Then $\delta = f''(\lambda_1)(z - \lambda_1)^2/2 + O(|z - \lambda_1|^3)$. The expansion for $\phi_1(\epsilon)$ follows easily from Theorem 3.1.

Next, we find the asymptotic expansion for ψ_1 assuming first that $f'(\lambda_1) \neq 0$. Let $\zeta_1 = f(\lambda_1)$ and $\zeta = f(z)$ for $z \in \Lambda_r(A, \lambda_1)$ for some small $r > 0$. The inverse function theorem states that the inverse of f is well defined near λ_1. Even though $f^{-1}(\zeta)$ in general is a set containing possibly several elements, we define $f^{-1}(\zeta)$ as the unique element in

$\Lambda_r(A,\lambda_1)$. Let $\delta = f^{-1}(\zeta) - f^{-1}(\zeta_1)$. By definition,

$$\begin{aligned}
\psi_1(s) &= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \inf\{r>0,\ \zeta\in f(\Lambda_r(A,\lambda_1))\}\\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \inf\{r>0,\ f^{-1}(\zeta)\in \Lambda_r(A,\lambda_1)\}\\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \inf\{r>0,\ \|(f^{-1}(\zeta)I-A)^{-1}\|\geq r^{-1}\}\\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \|(f^{-1}(\zeta)I-A)^{-1}\|^{-1}\\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \left\| Q\left(f^{-1}(\zeta)I - \begin{bmatrix}\lambda_1 I_m & \\ & D_2\end{bmatrix}\right)^{-1} Q^{-1}\right\|^{-1}\\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \left\| Q\begin{bmatrix}\delta^{-1} I_m & \\ & (f^{-1}(\zeta)I-D_2)^{-1}\end{bmatrix} Q^{-1}\right\|^{-1}\\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} |\delta| \left\| Q\begin{bmatrix} I_m & \\ & \delta(f^{-1}(\zeta)I-D_2)^{-1}\end{bmatrix} Q^{-1}\right\|^{-1}\\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \frac{|\delta|}{\|P\|} + O(|\delta|^2)\\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \frac{|\zeta-\zeta_1|}{|f'(\lambda_1)|\,\|P\|} + O(|\zeta-\zeta_1|^2)\\
&= \frac{s\,\|\tilde{P}\|}{|f'(\lambda_1)|\,\|P\|} + O(s^2).
\end{aligned}$$

In the above, we used the fact $\delta = \dfrac{\zeta-\zeta_1}{f'(\lambda_1)} + O(|\zeta-\zeta_1|^2)$ and Corollary 3.2. Now assume that $f'(\lambda_1)=0$ and $f''(\lambda_1)\neq 0$. Note that

$$\zeta-\zeta_1 = f(z)-f(\lambda_1) = \frac{f''(\lambda_1)(z-\lambda_1)^2}{2} + O(|z-\lambda_1|^3).$$

Given ζ in a small neighbourhood of $f(\lambda_1)$, there are two elements $z_\pm$ of $f^{-1}(\zeta)$ in a small neighbourhood of λ_1. They satisfy

$$|z_\pm - \lambda_1| = \frac{|\zeta-\zeta_1|^{1/2}}{\sqrt{|f''(\lambda_1)|/2}} + O(|\zeta-\zeta_1|). \tag{4.1}$$

Consequently,

$$\begin{aligned}
\psi_1(s) &= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \min\{\|(z_\pm I-A)^{-1}\|^{-1}\}\\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \min\left\{\frac{|\delta|}{\|P\|} + O(|\delta|^2),\ \delta = z_\pm-\lambda_1\right\}\\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \frac{|\zeta-\zeta_1|^{1/2}}{\sqrt{|f''(\lambda_1)|/2}\,\|P\|} + O(|\zeta-\zeta_1|)\\
&= \frac{s^{1/2}\,\|\tilde{P}\|^{1/2}}{\sqrt{|f''(\lambda_1)|/2}\,\|P\|} + O(s)
\end{aligned}$$

using (4.1) and Corollary 3.2. □

An immediate corollary of the above theorem is that

$$\phi_1(\psi_1(s)) = \begin{cases} s + O(s^2), & f'(\lambda_1) \neq 0; \\ s + O(s^{3/2}), & f'(\lambda_1) = 0,\ f''(\lambda_1) \neq 0, \end{cases} \tag{4.2}$$

and

$$\psi_1(\phi_1(\epsilon)) = \epsilon + O(\epsilon^2)$$

as long as $f'(\lambda_1)$ and $f''(\lambda_1)$ are not both zero.

THEOREM 4.2. *Let λ_1 be an eigenvalue of the matrix A of index $m \geq 2$, and let $A = QJQ^{-1}$ where J is a Jordan form of A of the form (3.4). Let f be a function which is analytic in some open set containing $\Lambda(A)$ satisfying $f'(\lambda_1) \neq 0$. Suppose $f(\lambda_1) \neq f(\lambda_j)$ for any other eigenvalue λ_j distinct from λ_1. For small $\epsilon > 0$,*

$$\phi_1(\epsilon) = |f'(\lambda_1)|\epsilon + O(\epsilon^{1+\frac{1}{m}}).$$

For small $s > 0$,

$$\psi_1(s) = \frac{s}{|f'(\lambda_1)|} + O(s^{1+\frac{1}{m}}).$$

Proof. Let $\delta = f(z) - f(\lambda_1) = f'(\lambda_1)(z - \lambda_1) + O(|z - \lambda_1|^2)$ for $z \in \Lambda_\epsilon(A, \lambda_1)$. Let N be as defined in (3.5). Then

$$\begin{aligned}
\phi_1(\epsilon) &= \sup_{z \in \Lambda_\epsilon(A,\lambda_1)} \|Q(f(z)I - f(J))^{-1}Q^{-1}\|^{-1} \\
&= \sup_{z \in \Lambda_\epsilon(A,\lambda_1)} \left\| Q \left[\begin{array}{cccc|c} \delta & -f'(\lambda_1) & \cdots & -\frac{f^{(m-1)}(\lambda_1)}{(m-1)!} & \\ & \ddots & \ddots & \vdots & \\ & & \delta & -f'(\lambda_1) & \\ & & & \delta & \\ \hline & & & & f(z)I - f(J_2) \end{array} \right]^{-1} Q^{-1} \right\|^{-1} \\
&= \sup_{z \in \Lambda_\epsilon(A,\lambda_1)} \frac{|\delta|^m}{|f'(\lambda_1)|^{m-1}\, \|N^{m-1}\|} + O(|\delta|^{m+1}) \\
&= \sup_{z \in \Lambda_\epsilon(A,\lambda_1)} \frac{|f'(\lambda_1)(z - \lambda_1)|^m}{|f'(\lambda_1)|^{m-1}\, \|N^{m-1}\|} + O(|z - \lambda_1|^{m+1}) \\
&= |f'(\lambda_1)|\, \epsilon + O(\epsilon^{1+\frac{1}{m}})
\end{aligned}$$

by (3.6) and Theorem 3.3.

Next, we find an asymptotic expansion for $\psi_1(s)$. Let $\delta = f^{-1}(\zeta) - f^{-1}(\zeta_1)$ where $\zeta \in \Lambda_s(f(A), \zeta_1)$ and $\zeta_1 = f(\lambda_1)$. Again, define $f^{-1}(\zeta)$ as the unique element in a small neighbourhood of λ_1. Now

ETNA
Kent State University
http://etna.math.kent.edu

$$
\begin{aligned}
\psi_1(s) &= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \|Q(f^{-1}(\zeta)I-J)^{-1}Q^{-1}\|^{-1} \\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \left\| Q \left[\begin{array}{cccc|c} \delta & -1 & & & \\ & \ddots & \ddots & & \\ & & \delta & -1 & \\ & & & \delta & \\ \hline & & & & f^{-1}(\zeta)I-J_2 \end{array}\right]^{-1} Q^{-1} \right\|^{-1} \\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \left\| Q \left[\begin{array}{cccc|c} \delta^{-1} & \delta^{-2} & \cdots & \delta^{-m} & \\ & \ddots & \ddots & \vdots & \\ & & \ddots & \delta^{-2} & \\ & & & \delta^{-1} & \\ \hline & & & & (f^{-1}(\zeta)I-J_2)^{-1} \end{array}\right] Q^{-1} \right\|^{-1} \\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \frac{|\delta|^m}{\|N^{m-1}\|} + O(|\delta|^{m+1}) \\
&= \sup_{\zeta\in\Lambda_s(f(A),\zeta_1)} \frac{|\zeta-\zeta_1|^m}{|f'(\lambda_1)|^m\,\|N^{m-1}\|} + O(|\zeta-\zeta_1|^{m+1}) \\
&= \frac{s}{|f'(\lambda_1)|} + O(s^{1+\frac{1}{m}})
\end{aligned}
$$

using Corollary 3.4. ☐

An immediate corollary of the above theorem is that

$$\psi_1(\phi_1(\epsilon)) = \epsilon + O(\epsilon^{1+\frac{1}{m}}) \text{ and } \phi_1(\psi_1(s)) = s + O(s^{1+\frac{1}{m}}). \tag{4.3}$$

Again for ease of exposition, we assumed that that there is only one Jordan block corresponding to λ_1 of size m. The result also holds in the general case of $k \geq 1$ such Jordan blocks. Using the facts discussed immediately following Corollary 3.4, a similar analysis also works for the other cases where $f(\lambda_1) = f(\lambda_j)$ for $j \neq 1$ or when $f'(\lambda_1) = 0$.

5. Eigenvalue perturbation theory for $f(A)$. We give an application of our pseudospectral mapping theorem for the eigenvalue perturbation problem of $f(A)$. Given a square matrix A, a non-constant function f analytic on an open set containing $\Lambda(A)$ and a positive ϵ, we wish to estimate how the eigenvalues of $f(A)$ change when A is perturbed by another matrix of norm at most ϵ. The relevant set is

$$\{\Lambda(f(A+E)),\ \|E\| \leq \epsilon\} = \{f(\Lambda(A+E)),\ \|E\| \leq \epsilon\} = f(\Lambda_\epsilon(A)).$$

Note the distinction between the above set and $\Lambda_\epsilon(f(A))$ which has already been estimated in Corollaries 3.2 and 3.4. Using (2.1) with $j = 1$ and $s = \psi_1^{-1}(\epsilon)$,

$$\Lambda_s(f(A), f(\lambda_1)) \subset f(\Lambda_\epsilon(A, \lambda_1)) \subset \Lambda_{\phi_1(\epsilon)}(f(A), f(\lambda_1)), \tag{5.1}$$

we obtain sharp lower and upper bounds on the size of the component containing a particular eigenvalue λ_1. (From the expansion of ψ_1 in Theorem 4.1 or 4.2 and the fact that f is non-constant, $\psi_1(s)$ is a strictly increasing function for $s \geq 0$ and small and so ψ_1^{-1} is uniquely defined.) Observe that in the setting of the previous theorems

$$\phi_1(\epsilon) = \phi_1(\psi_1(s)) = \begin{cases} s + O(s^2) \text{ or } s + O(s^{3/2}), & \text{in Theorem 4.1} \\ s + O(s^{1+\frac{1}{m}}), & \text{in Theorem 4.2} \end{cases} \tag{5.2}$$

ETNA
Kent State University
http://etna.math.kent.edu

by (4.2) and (4.3). Hence, the desired set $f(\Lambda_\epsilon(A,\lambda_1))$ is sandwiched between two sets of the same size to leading order in s.

Assuming the hypotheses of Theorem 4.1, we have from Corollary 3.2 and (5.1) that

$$\sup_{\zeta\in f(\Lambda_\epsilon(A,\lambda_1))} |\zeta - f(\lambda_1)| = \phi_1(\epsilon)\,\|\tilde{P}\| + O(\phi_1^2(\epsilon)).$$

From the expansion of ψ_1 in Theorem 4.1, we have

$$s = \psi_1^{-1}(\epsilon) = \begin{cases} |f'(\lambda_1)|\frac{\|P\|}{\|\tilde{P}\|}\,\epsilon + O(\epsilon^2), & f'(\lambda_1)\neq 0; \\ \frac{|f''(\lambda_1)|}{2}\,\frac{\|P\|^2}{\|\tilde{P}\|}\,\epsilon^2 + O(\epsilon^3), & f'(\lambda_1)=0,\ f''(\lambda_1)\neq 0. \end{cases}$$

We combine the above and (5.2) to obtain a sharp perturbation result for the eigenvalue $f(\lambda_1)$.

THEOREM 5.1. *Assume the hypotheses of Theorem 4.1. Then*

$$\sup_{\zeta\in f(\Lambda_\epsilon(A,\lambda_1))} |\zeta - f(\lambda_1)| = \begin{cases} |f'(\lambda_1)|\,\|P\|\,\epsilon + O(\epsilon^2), & f'(\lambda_1)\neq 0; \\ \frac{|f''(\lambda_1)|}{2}\,\|P\|^2\,\epsilon^2 + O(\epsilon^3), & f'(\lambda_1)=0,\ f''(\lambda_1)\neq 0. \end{cases}$$

In the literature, the leading order term of the right-hand side of the above is called the condition number of the eigenvalue $f(\lambda_1)$ when A is subject to a perturbation of size at most ϵ. It is interesting that this condition number is independent of any information about another eigenvalue λ_k in case $f(\lambda_k)=f(\lambda_1)$. The next term in the expansion can be interpreted as the departure from non-normality of the eigenvalue.

In the same way, we have

THEOREM 5.2. *Assume the hypotheses of Theorem 4.2. Then*

$$\sup_{\zeta\in f(\Lambda_\epsilon(A,\lambda_1))} |\zeta - f(\lambda_1)| = |f'(\lambda_1)|\,\|N^{m-1}\|^{1/m}\,\epsilon^{1/m} + O(\epsilon^{2/m}).$$

6. Examples and numerical results. In this section, we work out two examples analytically and supply three numerical experiments to confirm the theoretical estimates for the functions ϕ_j and ψ_j as well as two numerical experiments for the eigenvalue perturbation problem.

EXAMPLE 6.1. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. This matrix is normal and everything can be worked out analytically. The eigenvalues are $\lambda_1 = 1,\ \lambda_2=-1$. Take $f(z)=z^2+2z$. First consider $\lambda_1 = 1$ and observe that $f'(1)=4$.

$$\begin{aligned} \phi_1(\epsilon) &= \sup_{z\in\Lambda_\epsilon(A,1)} \|(f(z)I - f(A))^{-1}\|^{-1} \\ &= \sup_{z\in\Lambda_\epsilon(A,1)} \left\| \begin{bmatrix} z^2+2z-3 & \\ & z^2+2z+1 \end{bmatrix}^{-1} \right\|^{-1} \\ &= \sup_{|z-1|\le\epsilon} |z^2+2z-3| \\ &= \sup_{|\zeta|\le\epsilon} |\zeta^2+4\zeta| \\ &= 4\epsilon+\epsilon^2. \end{aligned}$$

Next,

$$\begin{aligned}\psi_1(s) &= \sup_{z\in\Lambda_s(f(A),3)} \|(f^{-1}(z)I-A)^{-1}\|^{-1}\\ &= \sup_{|z-3|\le s}\left\|\begin{bmatrix}\sqrt{1+z}-2 & \\ & \sqrt{1+z}\end{bmatrix}^{-1}\right\|^{-1}\\ &= \sup_{|z-3|\le s}|\sqrt{1+z}-2|\\ &= \sup_{|\zeta|\le s}|2\sqrt{1+\zeta/4}-2|\\ &= 2-2\sqrt{1-s/4}=\frac{s}{4}+\frac{s^2}{64}+\cdots.\end{aligned}$$

Next, consider the eigenvalue $\lambda_2=-1$ where $f'(-1)=0$ and $f''(-1)=2$. We find

$$\begin{aligned}\phi_2(\epsilon) &= \sup_{z\in\Lambda_\epsilon(A,-1)}\left\|\begin{bmatrix}z^2+2z-3 & \\ & z^2+2z+1\end{bmatrix}^{-1}\right\|^{-1}\\ &= \sup_{|z+1|\le\epsilon}|z^2+2z+1|\\ &= \epsilon^2\end{aligned}$$

and

$$\begin{aligned}\psi_2(s) &= \sup_{z\in\Lambda_s(f(A),-1)}\left\|\begin{bmatrix}\sqrt{1+z}-2 & \\ & \sqrt{1+z}\end{bmatrix}^{-1}\right\|^{-1}\\ &= \sup_{|z+1|\le s}|1+z|^{1/2}\\ &= s^{1/2}.\end{aligned}$$

These results agree with Theorem 4.1.

EXAMPLE 6.2. Let $A=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$. The eigenvalue $\lambda_1=0$ has index $m=2$. Take $f(z)=z^2+z$. Observe that $f'(0)=1$ and $f(A)=A$. Now

$$\begin{aligned}\phi_1(\epsilon) &= \sup_{z\in\Lambda_\epsilon(A,0)}\|(f(z)I-f(A))^{-1}\|^{-1}\\ &= \sup_{z\in\Lambda_\epsilon(A,0)}\left\|\begin{bmatrix}f(z)^{-1} & f(z)^{-2}\\ & f(z)^{-1}\end{bmatrix}\right\|^{-1}\\ &= \sup_{|z|\le\sqrt{\epsilon+\epsilon^2}}|z|^2|1+z|^2+\cdots\\ &= \epsilon+2\epsilon^{3/2}+\cdots.\end{aligned}$$

In the calculation of ψ_1 below, let $\zeta=f^{-1}(z)=(-1+\sqrt{1+4z})/2\approx z-z^2$ for a small z.

TABLE 6.1

f_1	$\phi_1(\epsilon)$	$\psi_1(\epsilon)$
$a=1,\epsilon=10^{-3}$	2.0002×10^{-3} (2×10^{-3})	5.0051×10^{-4} (5×10^{-4})
$a=1,\epsilon=10^{-4}$	2.0000×10^{-4} (2×10^{-4})	5.0005×10^{-5} (5×10^{-5})
$a=10,\epsilon=10^{-4}$	2.0005×10^{-4} (2×10^{-4})	5.0053×10^{-5} (5×10^{-5})

f_2	$\phi_1(\epsilon)$	$\psi_1(\epsilon)$
$a=1,\epsilon=10^{-3}$	1.7321×10^{-6} (1.7321×10^{-6})	2.4037×10^{-2} (2.4028×10^{-2})
$a=1,\epsilon=10^{-4}$	1.7321×10^{-8} (1.7321×10^{-8})	7.5986×10^{-3} (7.5984×10^{-3})
$a=10,\epsilon=10^{-4}$	1.4177×10^{-7} (1.4177×10^{-7})	2.6576×10^{-3} (2.6558×10^{-3})

Then

$$\begin{aligned}\psi_1(s) &= \sup_{z\in\Lambda_s(f(A),0)} \|(\zeta I - A)^{-1}\|^{-1}\\ &= \sup_{z\in\Lambda_s(A,0)} \left\|\begin{bmatrix}\zeta^{-1} & \zeta^{-2}\\ & \zeta^{-1}\end{bmatrix}\right\|^{-1}\\ &= \sup_{|z|\le\sqrt{s+s^2}} |z|^2|1-z|^2+\cdots\\ &= s+2s^{3/2}+\cdots.\end{aligned}$$

This example illustrates the correctness of Theorem 4.2.

EXAMPLE 6.3. Let a be any real number and

$$A=\begin{bmatrix}0 & 0 & a\\ & 0 & a\\ & & 1\end{bmatrix}. \tag{6.1}$$

Note that A is diagonalizable with an eigenvalue $\lambda_1=0$ of multiplicity two. A calculation leads to $A=QDQ^{-1}$ where

$$Q=\begin{bmatrix}1 & & a\\ & 1 & a\\ & & 1\end{bmatrix},\qquad D=\begin{bmatrix}0 & & \\ & 0 & \\ & & 1\end{bmatrix}.$$

The projection onto the eigenspace corresponding to the eigenvalue 0 is

$$P=Q\begin{bmatrix}1 & & \\ & 1 & \\ & & 0\end{bmatrix}Q^{-1}=\begin{bmatrix}1 & & -a\\ & 1 & -a\\ & & 0\end{bmatrix}.$$

We find that $\|P\|=\sqrt{2a^2+1}$. Results of numerical computations for the eigenvalue 0 and $f_1(z)=z^2+2z$ and $f_2(z)=z^2$ are shown in Table 6.1. Note that $f_1'(0)=2,\ f_2'(0)=0$, $f_2''(0)=2$ and $\tilde{P}=P$ for both functions. In the table, the numbers in parentheses denote the values predicted by first order expansions in Theorem 4.1: $\phi_1(\epsilon)\approx 2\epsilon,\ \psi_1(s)\approx s/2$ for f_1 and $\phi_1(\epsilon)\approx\|P\|\,\epsilon^2,\ \psi_1(s)\approx\sqrt{s}/\|P\|^{1/2}$ for f_2.

EXAMPLE 6.4. Consider the same matrix as in the above example. Take $f_3(z)=z^2-z$ and $f_4(z)=z^4-z^2$. Observe that $f_3(0)=0=f_3(1),\ f_3'(0)=-1$ and $f_4(0)=0=f_4(1)$, $f_4'(0)=0,\ f_4''(0)=-2$. For both f_3 and f_4, we have $\tilde{P}=I$. Table 6.2 shows the results of

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 6.2

f_3	$\phi_1(\epsilon)$	$\psi_1(\epsilon)$
$a=1, \epsilon=10^{-3}$	1.7351×10^{-3} (1.7321×10^{-3})	5.7754×10^{-4} (5.7735×10^{-4})
$a=1, \epsilon=10^{-4}$	1.7324×10^{-4} (1.7321×10^{-4})	5.7737×10^{-5} (5.7735×10^{-5})
$a=10, \epsilon=10^{-4}$	1.4198×10^{-3} (1.4177×10^{-3})	7.0535×10^{-6} (7.0535×10^{-6})

f_4	$\phi_1(\epsilon)$	$\psi_1(\epsilon)$
$a=1, \epsilon=10^{-3}$	3.0000×10^{-6} (3×10^{-6})	1.8252×10^{-2} (1.8257×10^{-2})
$a=1, \epsilon=10^{-4}$	3.0000×10^{-8} (3×10^{-8})	5.7733×10^{-3} (5.7735×10^{-3})
$a=10, \epsilon=10^{-4}$	2.0100×10^{-6} (2.01×10^{-6})	7.0533×10^{-4} (7.0535×10^{-4})

TABLE 6.3

f_5	$\phi_1(\epsilon)$	$\psi_1(\epsilon)$
$a=1, \epsilon=10^{-3}$	2.0333×10^{-3} (2×10^{-3})	5.2025×10^{-4} (5×10^{-4})
$a=1, \epsilon=10^{-4}$	2.0115×10^{-4} (2×10^{-4})	5.0634×10^{-5} (5×10^{-5})
$a=10, \epsilon=10^{-4}$	2.0183×10^{-4} (2×10^{-4})	5.7814×10^{-5} (5×10^{-5})

numerical computations for the eigenvalue $\lambda_1 = 0$. They agree with the first order expansions of Theorem 4.1: $\phi_1(\epsilon) \approx \|P\|\epsilon,\ \psi_1(s) \approx s/\|P\|$ for f_3 and $\phi_1(\epsilon) \approx \|P\|^2\,\epsilon^2,\ \psi_1(s) \approx \sqrt{s}/\|P\|$ for f_4.

EXAMPLE 6.5. Let a be any real number and

$$A = \begin{bmatrix} 0 & a & \\ & 0 & a \\ & & 1 \end{bmatrix}. \tag{6.2}$$

Note that A is not diagonalizable with an eigenvalue $\lambda_1 = 0$ of algebraic multiplicity two and geometric multiplicity one. A calculation leads to $A = QJQ^{-1}$ where

$$Q = \begin{bmatrix} 1 & 0 & a \\ & 1/a & 1 \\ & 0 & 1/a \end{bmatrix}, \qquad J = \begin{bmatrix} 0 & 1 & \\ & 0 & \\ & & 1 \end{bmatrix}.$$

The nilpotent matrix corresponding to the eigenvalue 0 is

$$N = \begin{bmatrix} 0 & a & -a^2 \\ & 0 & \\ & & 0 \end{bmatrix}.$$

We find that $\|N\| = \sqrt{a^2+a^4}$. Define $f_5(z) = z^2 + 2z$ and note that $f_5'(0) = 2$. The numerical results are shown in Table 6.3. They agree with the first order expansions predicted by Theorem 4.2: $\phi_1(\epsilon) \approx 2\epsilon,\ \psi_1(s) \approx s/2$.

EXAMPLE 6.6. Take the matrix in (6.1) with $a = 10$ and $f(z) = z^3 - z$. Note that $f(0) = f(1) = 0$ and $f'(0) = -1,\ f'(1) = 2$. With $\epsilon = 10^{-3}$, the curve $f(\partial\Lambda_\epsilon(A, \lambda_j))$ and its approximation by a disk of radius given by the first term of the expansion given in Theorem 5.1 are shown in Figure 6.1. For $f(z) = z^4 - z^2$ where $f(0) = 0 = f(1)$ and $f'(0) = 0,\ f''(0) = -2$, see Figure 6.2. In both instances, there is excellent agreement with the theoretical estimate, demonstrating that indeed the leading order behaviour is determined

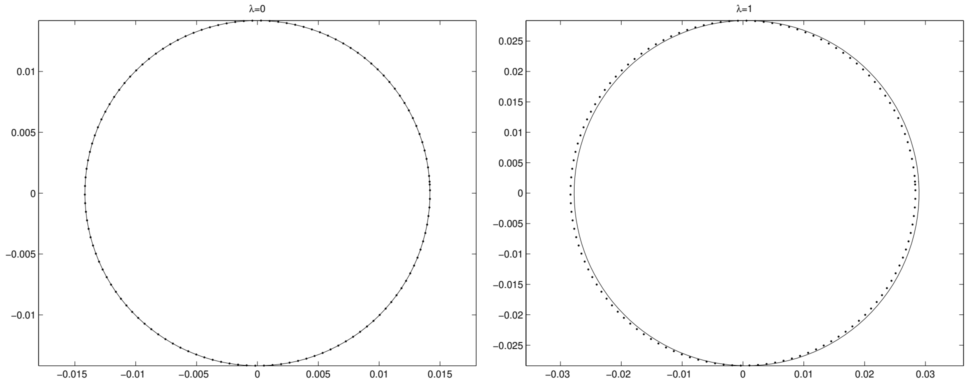

FIG. 6.1. $f(z) = z^3 - z$ *and* A *given by* (6.1). *The dotted curve denotes the circle given by the first term of the expansion in Theorem 5.1 while the solid curve denotes* $f(\partial\Lambda_\epsilon(A, \lambda_j))$.

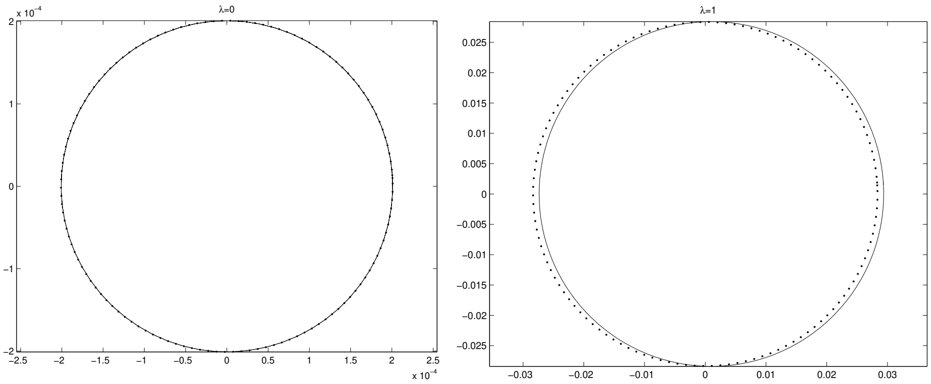

FIG. 6.2. $f(z) = z^4 - z^2$ *and* A *given by* (6.1). *The dotted curve denotes the circle given by the first term of the expansion in Theorem 5.1 while the solid curve denotes* $f(\partial\Lambda_\epsilon(A, \lambda_j))$.

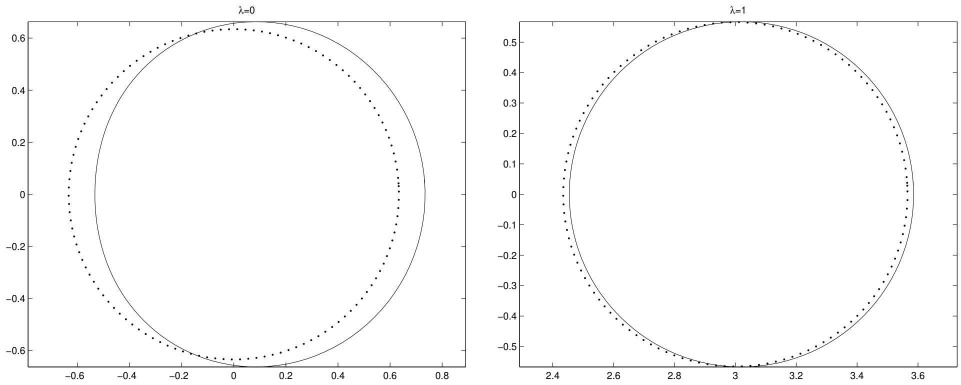

FIG. 6.3. $f(z) = z^2 + 2z$ *and* A *given by* (6.2). *The dotted curve denotes the circle given by the first term of the expansion in Theorem 5.2 while the solid curve denotes* $f(\partial\Lambda_\epsilon(A, \lambda_j))$.

by the eigenvalue λ_j in question and independent of the other eigenvalues λ_k for which $f(\lambda_k) = f(\lambda_j)$.

EXAMPLE 6.7. Take the matrix in (6.2) with $a = 10$ and $f(z) = z^2+2z$. With $\epsilon = 10^{-3}$, the curve $f(\partial\Lambda_\epsilon(A, \lambda_j))$ and its approximation by a disk of radius given by the first term of the expansion given in Theorem 5.2 are shown in Figure 6.3. The discrepancy between the actual and predicted curves is more significant for $\lambda = 0$. This can be attributed to the large value of $\|N\| \approx O(a^2)$ and the fact that the error behaves like $O(\epsilon)$. The discrepancy decreases as the value of a decreases or as ϵ decreases.

Acknowledgment. I am grateful to two anonymous referees for their careful reading of the manuscript and suggestions which have improved the presentation.

REFERENCES

[1] 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.
[2] ———, *Spectral conditioning and pseudospectral growth*, Numer. Math., 107 (2007), pp. 27–37.
[3] M. KAROW, *Eigenvalue condition numbers and a formula of Burke, Lewis and Overton*, Electron. J. Linear Algebra, 15 (2006), pp. 143–153.
[4] T. KATO, *A Short Introduction to Perturbation Theory for Linear Operators*, Springer, New York, 1982.
[5] S. H. LUI, *A pseudospectral mapping theorem*, Math. Comp., 72 (2003), pp. 1841–1854.
[6] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra*, Princeton University Press, Princeton, 2005.
[7] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, UK, 1965.