Eigenvalue condition numbers and a formula of Burke, Lewis and Overton

Michael Karow
Berlin University of Technology
karow@math.TU-Berlin.de

October 18, 2004

Abstract

In [1] a first order expansion has been given for the minimum singular value of $A - zI$, $z \in \mathbb{C}$, about a nonderogatory eigenvalue λ of $A \in \mathbb{C}^{n\times n}$. This note investigates the relationship of the expansion with the Jordan canonical form of A. Furthermore, formulas for the condition number of eigenvalues are derived from the expansion.

1 Introduction

By $\pi_\Sigma(A)$ we denote the product of the nonzero singular values of the matrix $A\in\mathbb{C}^{n\times m}$, counting multiplicities. If A is square then $\Lambda(A)$ denotes the spectrum and $\pi_\Lambda(A)$ stands for the product of the nonzero eigenvalues, counting multiplicities. The subject of this note is the ratio

$$q(A,\lambda) := \frac{\pi_\Sigma(A-\lambda\, I_n)}{|\pi_\Lambda(A-\lambda\, I_n)|}, \quad \lambda\in\Lambda(A).$$

In [1] the following first order expansion has been given for the function

$$z \mapsto \sigma_{\min}(A - zI_n),\ z\in\mathbb{C},$$

where $\sigma_{\min}(\cdot)$ denotes the minimum singular value and I_n is the $n\times n$ identity matrix.

Theorem 1.1 *Let $\lambda \in \mathbb{C}$ be a nonderogatory eigenvalue of algebraic multiplicity m of the matrix $A\in\mathbb{C}^{n\times n}$. Then*

$$\sigma_{\min}(A - zI_n) = \frac{|z-\lambda|^m}{q(A,\lambda)} + \mathcal{O}(|z-\lambda|^{m+1}), \quad z\in\mathbb{C}.$$

The relevance of this result for the perturbation theory of eigenvalues is as follows. The $\epsilon-$ pseudospectrum of $A\in\mathbb{C}^{n\times n}$ with respect to the spectral norm, $\|\cdot\|$, is defined by

$$\Lambda_\epsilon(A) = \{\ z\in\mathbb{C} \mid\ z\in\Lambda(A+\Delta),\ \Delta\in\mathbb{C}^{n\times n},\ \|\Delta\| \le \epsilon\ \}.$$

In words, $\Lambda_\epsilon(A)$ is the set of all eigenvalues of all matrices of the form $A + \Delta$ where the spectral norm of the perturbation Δ is bounded by $\epsilon > 0$. It is well known that

$$\Lambda_\epsilon(A) = \{ \; z \in \mathbb{C} \mid \; \sigma_{\min}(A - zI) \leq \epsilon \; \}.$$

Theorem 1.1 yields an estimate for the size of pseudospectra for small ϵ: Roughly speaking if ϵ is small enough then the connected component of $\Lambda_\epsilon(A)$ that contains the eigenvalue λ is approximately a disk of radius $(\, q(A,\lambda)\, \epsilon\,)^{1/m}$ about λ. It follows that $q(A,\lambda)^{1/m}$ is the Hölder condition number of λ. We discuss this in detail in Section 4.

However, the main concern of this note is to establish the relationship of $q(A,\lambda)$ with the Jordan decomposition of A. For a simple eigenvalue the relationship is as follows. Let $x, y \in \mathbb{C}^n$ be a right and a left eigenvector of A to the eigenvalue λ respectively. Then $P = (y^*x)^{-1}xy^*$ is the projection onto the one dimensional eigenspace $\mathbb{C}\, x$. The kernel of P is the direct sum of all generalized eigenspaces belonging to the eigenvalues different from λ. As is well known the condition number of λ equals the norm of P. Combined with the considerations above this yields that

$$q(A,\lambda) = \|P\|. \tag{1}$$

In Section 3 we give an elementary proof of the identity (1) without using Theorem 1.1. Furthermore, we show that for a nondegoratory eigenvalue of algebraic multiplicity $m \geq 2$,

$$q(A,\lambda) = \|N^{m-1}\|, \tag{2}$$

where N is the nilpotent operator associated with λ in the Jordan decomposition of A. The formulas (1) and (2) are the main results of this note. The proofs also show that the assumption that λ is nonderogatory is necessary.

The next section contains some prelimaries about the computation of the products $\pi_\Sigma(A)$ and $\pi_\Lambda(A)$ and about the relationship of the Schur form of A with the Jordan decomposition.

Throughout this note, $\|\cdot\|$ stands for the spectral norm.

2 Preliminaries

Below we list some easily verified properties of $\pi_\Lambda(A)$, the product of the nonzero eigenvalues of A, and of $\pi_\Sigma(A)$, the product of the nonzero singular values of A. In the sequel A^T and A^* denote the transpose and the conjugate transpose of A respectively.

(*a*) If $A \in \mathbb{C}^{n\times n}$ is nonsingular then $\pi_\Lambda(A) = \det(A)$.

(*b*) For any $A \in \mathbb{C}^{n\times n}$: $\pi_\Lambda(A^T) = \pi_\Lambda(A)$ and $\pi_\Lambda(A^*) = \overline{\pi_\Lambda(A)}$.

(*c*) Let $S \in \mathbb{C}^{n\times n}$ be nonsingular. Then for any $A \in \mathbb{C}^{n\times n}$, $\pi_\Lambda(SAS^{-1}) = \pi_\Lambda(A)$.

(*d*) Let $A_{11} \in \mathbb{C}^{n\times n}$, $A_{22} \in \mathbb{C}^{m\times m}$ and $A_{12} \in \mathbb{C}^{n\times m}$. Then

$$\pi_\Lambda\left(\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}\right) = \pi_\Lambda(A_{11})\, \pi_\Lambda(A_{22}).$$

(e) For any $A \in \mathbb{C}^{n\times m}$, $\pi_\Sigma(A)^2 = \pi_\Lambda(A^*A) = \pi_\Lambda(AA^*)$.

(f) If $A \in \mathbb{C}^{n\times n}$ is nonsingular then $\pi_\Sigma(A) = |\det(A)| = |\pi_\Lambda(A)|$.

(g) Let $U \in \mathbb{C}^{n\times n}$ and $V \in \mathbb{C}^{m\times m}$ be unitary. Then for any $A \in \mathbb{C}^{n\times m}$, $\pi_\Sigma(UAV) = \pi_\Sigma(A)$.

In the next section we need the lemmas below.

Lemma 2.1 *Let $M \in \mathbb{C}^{n\times n}$ be nonsingular, $X \in \mathbb{C}^{p\times n}$ and $Y = XM^{-1}$. Then*

$$\pi_\Sigma\left(\begin{bmatrix} M \\ X \end{bmatrix}\right) = \pi_\Sigma(M)\sqrt{\det(I + Y^*Y)}.$$

Proof: We have

$$\begin{aligned}
\pi_\Sigma\left(\begin{bmatrix} M \\ X \end{bmatrix}\right)^2 &= \pi_\Lambda\left(\begin{bmatrix} M^* & X^* \end{bmatrix}\begin{bmatrix} M \\ X \end{bmatrix}\right) \\
&= \det(M^*M + X^*X) \\
&= \det(M^*(I + Y^*Y)M) \\
&= \det(M^*)\det(M)\det(I + Y^*Y) \\
&= \pi_\Sigma(M)^2 \det(I + Y^*Y).
\end{aligned}$$

□

Lemma 2.2 *Let $Y \in \mathbb{C}^{n\times m}$. Then $\|I_n + Y^*Y\| = \|I_m + YY^*\|$ and $\det(I_n + Y^*Y) = \det(I_m + YY^*)$.*

Proof: The case $Y = 0$ is trivial. Let $Y \neq 0$. The matrices Y and Y^* have the same nonzero singular values $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_p$ say. The eigenvalues different from 1 of both $I_n + Y^*Y$ and $I_m + YY^*$ are $1+\sigma_1^2 \geq 1+\sigma_2^2 \ldots \geq 1+\sigma_p^2$. Thus $\|I_n + Y^*Y\| = \|I_m + YY^*\| = 1 + \sigma_1^2$ and $\det(I_n + Y^*Y) = \det(I_m + YY^*) = \prod_{k=1}^p (1 + \sigma_k^2)$. □

We proceed with remarks on the Jordan decomposition. Let $\lambda_1, \ldots, \lambda_\kappa$ be the pairwise different eigenvalues of $A \in \mathbb{C}^{n\times n}$. Let $\mathcal{X}_j = \ker(A - \lambda_j I_n)^n$ be the generalized eigenspaces. By the Jordan decomposition theorem we have

$$A = \sum_{j=1}^{\kappa} (\, \lambda_j P_j + N_j \,), \tag{3}$$

where $P_1, \ldots, P_\kappa \in \mathbb{C}^{n\times n}$ are the projectors of direct decomposition $\mathbb{C}^n = \bigoplus_{j=1}^{\kappa} \mathcal{X}_j$, i.e.

$$P_j^2 = P_j, \quad \text{range}(P_j) = \mathcal{X}_j, \quad \ker(P_j) = \bigoplus_{k=1, k\neq j}^{\kappa} \mathcal{X}_k,$$

and $N_1, \ldots, N_\kappa \in \mathbb{C}^{n\times n}$ are the nilpotent matrices $N_j = (A - \lambda_j I_n)P_j$. The eigenvalue λ_j is said to be

- semisimple (nondefective) if $\mathcal{X}_j = \ker(A - \lambda_j I_n)$,
- simple if $\dim \mathcal{X}_j = 1$,
- nonderogatory if $\dim \ker(A - \lambda_j I_n) = 1$.

In the following m denotes the algebraic multiplicity of λ_j. Note that if $m \geq 2$ then λ_j is nonderogatory if and only if $N_j^{m-1} \neq 0$. We now recall how to obtain the operators P_j and N_j from a Schur form of A. We only consider the nontrivial case that A has at least two different eigenvalues. By the Schur decomposition theorem there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that

$$U^* A U = \begin{bmatrix} \lambda_j I_m + T & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

where $A_{12} \in \mathbb{C}^{m\times(n-m)}$, $A_{22} \in \mathbb{C}^{(n-m)\times(n-m)}$, $\Lambda(A_{22}) = \Lambda(A) \setminus \{\lambda_j\}$ and $T \in \mathbb{C}^{n\times n}$ is strictly upper triangular,

$$T = \begin{bmatrix} 0 & t_{12} & \dots & \dots & t_{1m} \\ & \ddots & t_{13} & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & t_{m-1,m} \\ & & & & 0 \end{bmatrix}.$$

If $m = 1$ (i.e. λ_j is simple) then T is the 1×1 zero matrix. Since the spectra of T and $A_{22} - \lambda_j I_{n-m}$ are disjoint the Sylvester equation

$$R(A_{22} - \lambda_j I_{n-m}) - TR = A_{12}. \tag{4}$$

has a unique solution $R \in \mathbb{C}^{m\times(n-m)}$.

Proposition 2.3 *With the notation above the projector onto the generalized eigenspace and the nilpotent operator associated with λ_j are given by*

$$P_j = U \begin{bmatrix} I_m & -R \\ 0 & 0 \end{bmatrix} U^*, \qquad \text{and} \qquad N_j = U \begin{bmatrix} T & -TR \\ 0 & 0 \end{bmatrix} U^*.$$

For any integer $\ell \geq 1$ we have

$$N_j^\ell = U \begin{bmatrix} T^\ell & -T^\ell R \\ 0 & 0 \end{bmatrix} U^*. \tag{5}$$

The spectral norms of P_j and of N_j^ℓ satisfy

$$\|P_j\| = \|I_m + RR^*\|^{1/2} \tag{6}$$
$$\|N_j^\ell\| = \|T^\ell (I_m + RR^*)(T^*)^\ell\|^{1/2}. \tag{7}$$

Proof: Let $X_1 := U \begin{bmatrix} I_m \\ 0 \end{bmatrix} \in \mathbb{C}^{n\times m}$, $X_2 := U \begin{bmatrix} R \\ I_{n-m} \end{bmatrix} \in \mathbb{C}^{n\times(n-m)}$. Then $A X_1 = X_1 (\lambda_j I_m + T)$ and (4) yields that $A X_2 = X_2 A_{22}$. Hence, range(X_1) and range(X_2) are complementary invariant subspaces of A. Furthermore it follows that for any $\lambda \in \mathbb{C}$ and any integer $\ell \geq 1$,

$$(A - \lambda I_n)^\ell X_1 = X_1 ((\lambda_j - \lambda) I_m + T)^\ell, \qquad (A - \lambda I_n)^\ell X_2 = X_2 (A_{22} - \lambda I_{n-m})^\ell. \tag{8}$$

Let $A_{22} = \sum_{k \leq \kappa, k \neq j} (\lambda_k \widehat{P}_k + \widehat{N}_k)$ be the Jordan decomposition of A_{22}. Set

$$\begin{aligned} P_k &:= U \begin{bmatrix} 0 & R\widehat{P}_k \\ 0 & \widehat{P}_k \end{bmatrix} U^*, \qquad k = 1, \dots, \kappa, \quad k \neq j, \\ P_j &:= U \begin{bmatrix} I & -R \\ 0 & 0 \end{bmatrix} U^*, \\ N_k &:= (A - \lambda_k I) P_k, \qquad k = 1, \dots, \kappa. \end{aligned} \tag{9}$$

Using the relations (8) it is staightforward to verify that $A = \sum_{k=1}^{\kappa} (\lambda_k P_k + N_k)$ is the Jordan decomposition of A. The formulas (8), (6) and (7) are immediate from (9). □

We give an expression for $\|N_j^{m-1}\|$ which is a bit more explicite than formula (7). First note that if λ_j has algebraic multiplicity $m \geq 2$ then

$$T^{m-1} = \begin{bmatrix} 0 & \dots & 0 & \tau \\ \vdots & & \vdots & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix}, \quad \text{where} \quad \tau = \prod_{k=1}^{m-1} t_{k,k+1}.$$

Let $e_m^T = [0 \dots 0 \ 1]^T \in \mathbb{C}^m$ and $r = e_m^T R$. Then r is the lower row of R. Since the lower row of TR is zero it follows from the Sylvester equation (4) that

$$r = e_m^T A_{12} (A_{22} - \lambda_j I_m)^{-1}. \tag{10}$$

From (5) or (7) we obtain

Proposition 2.4 *Suppose λ_j has algebraic multiplicity $m \in \{2, \dots, n-1\}$. Then*

$$\|N_j^{m-1}\| = |\tau| \sqrt{1 + \|r\|^2}.$$

3 Main result

We are now in a position to state and proof our main result on the ratio

$$q(A, \lambda_j) = \frac{\pi_\Sigma(A - \lambda_j I_n)}{|\pi_\Lambda(A - \lambda_j I_n)|}, \qquad \lambda_j \in \Lambda(A). \tag{11}$$

Theorem 3.1 *Let $\lambda_j \in \mathbb{C}$ be an eigenvalue of $A \in \mathbb{C}^{n \times n}$. Let P_j and N_j be the eigenprojector and the nilpotent operator associated with λ_j. Then the following holds.*

(*a*) *If* λ_j *is a semisimple eigenvalue then* $q(A,\lambda_j)=\pi_\Sigma(P_j)$.

(*b*) *If* λ_j *is a simple eigenvalue then* $q(A,\lambda_j)=\|P_j\|$.

(*c*) *If* λ_j *is a nonderogatory eigenvalue of algebraic multiplicity* $m\geq 2$ *then*

$$q(A,\lambda_j)=\|N_j^{m-1}\|.$$

Proof: First, we treat the case that A has at least two different eigenvalues. In view of Proposition 2.3 and since the products $\pi_\Sigma(A-\lambda_j I_n)$, $\pi_\Lambda(A-\lambda_j I_n)$ are invariant under unitary similarity transformations we may assume that

$$A=\begin{bmatrix}\lambda_j I_m+T & A_{12}\\ 0 & A_{22}\end{bmatrix},\quad P_j=\begin{bmatrix}I_m & -R\\ 0 & 0\end{bmatrix},$$

where $\Lambda(A_{22})=\Lambda(A)\setminus\{\lambda_j\}$, $T\in\mathbb{C}^{n\times n}$ is strictly upper triangular and $R\in\mathbb{C}^{m\times(n-m)}$ is the solution of the Sylvester equation $R(A_{22}-\lambda_j I_{n-m})-TR=A_{12}$.
(*a*). Suppose λ_j is semisimple. Then $T=0$ and $R(A_{22}-\lambda_j I_{n-m})=A_{12}$. Thus,

$$\begin{aligned}(A-\lambda_j I_n)^*(A-\lambda_j I_n) &= \begin{bmatrix}0 & 0\\ 0 & (A_{22}-\lambda_j I_{n-m})^*(A_{22}-\lambda_j I_{n-m})+A_{12}^*A_{12}\end{bmatrix}\\ &= \begin{bmatrix}0 & 0\\ 0 & (A_{22}-\lambda_j I_{n-m})^*(I_{n-m}+R^*R)(A_{22}-\lambda_j I_{n-m})\end{bmatrix}.\end{aligned}$$

Thus

$$\begin{aligned}\pi_\Sigma(A-\lambda_j I_n)^2 &= \det((A_{22}-\lambda_j I_{n-m})^*(I_{n-m}+R^*R)(A_{22}-\lambda_j I_{n-m}))\\ &= |\det(A_{22}-\lambda_j I_{n-m})|^2\det(I_{n-m}+R^*R)\\ &= |\pi_\Lambda(A_n-\lambda_j I)|^2\det(I_{n-m}+R^*R). \end{aligned}\tag{12}$$

Furthermore we have $P_jP_j^*=\begin{bmatrix}0 & I_m+RR^*\\ 0 & 0\end{bmatrix}$ and hence

$$\pi_\Sigma(P_j)^2=\det(I_m+RR^*)=\det(I_{n-m}+R^*R).\tag{13}$$

The latter equation holds by Lemma 2.2. By combining (12) and (13) we obtain (*a*).
(*b*). If $m=1$ then P_j has rank 1 and hence, $\pi_\Sigma(P_j)=\|P_j\|$. Thus (*b*) follows from (*a*).
(*c*) Suppose $m\geq 2$ and λ_j is nonderogatory. Then $T=\begin{bmatrix}0 & D\\ \vdots & \\ 0\ \dots & 0\end{bmatrix}$, where $D\in\mathbb{C}^{(m-1)\times(m-1)}$ is upper triangular and nonsingular. In the following we write $A_{12}=\begin{bmatrix}\tilde{A}\\ a\end{bmatrix}$, where a is the lower row of A_{12}. Let r denote the lower row of R. By Formula (10) we have

$$r=a(A_{22}-\lambda_j I)^{-1}.\tag{14}$$

Let us determine $\pi_\Sigma(A)$. Since removing of a column of zeros and a permutation of rows does not change the nonzero singular values of a matrix we have

$$\pi_\Sigma(A-\lambda_j I_n) \;=\; \pi_\Sigma\left(\begin{bmatrix} \begin{smallmatrix}0\\ \vdots\end{smallmatrix} & D & \tilde{A} \\ \begin{smallmatrix}0 & \dots & 0\end{smallmatrix} & & a \\ 0 & & A_{22}-\lambda_j I_{n-m}\end{bmatrix}\right) = \pi_\Sigma\left(\begin{bmatrix} D & \tilde{A} \\ 0 & A_{22}-\lambda_j I_{n-m} \\ \begin{smallmatrix}0\dots 0\end{smallmatrix} & a\end{bmatrix}\right).$$

Lemma 2.1 yields

$$\begin{aligned}\pi_\Sigma\left(\begin{bmatrix} D & \tilde{A} \\ 0 & A_{22}-\lambda_j I_{n-m} \\ \begin{smallmatrix}0\dots 0\end{smallmatrix} & a\end{bmatrix}\right) &= \pi_\Sigma\left(\begin{bmatrix} D & \tilde{A} \\ 0 & A_{22}-\lambda_j I\end{bmatrix}\right)\sqrt{\det(1+yy^*)} \\ &= |\det(D)\det(A_{22}-\lambda_j I)|\,\sqrt{1+\|y\|^2} \\ &= |\pi_\Lambda(A-\lambda_j I)|\;|\det(D)|\,\sqrt{1+\|y\|^2}\end{aligned}$$

where

$$y = \begin{bmatrix}\begin{smallmatrix}0\dots 0\end{smallmatrix} & a\end{bmatrix}\begin{bmatrix} D & \tilde{A} \\ 0 & A_{22}-\lambda_j I\end{bmatrix}^{-1}.$$

From (14) it follows that $y = \begin{bmatrix}\begin{smallmatrix}0\dots 0\end{smallmatrix} & r\end{bmatrix}$ and hence, $\|y\| = \|r\|$. In summary,

$$\pi_\Sigma(A-\lambda_j I_n) \;=\; |\pi_\Lambda(A-\lambda_j I_n)|\,|\det(D)|\,\sqrt{1+\|r\|^2}.$$

But $|\det(D)|\sqrt{1+\|r\|^2} = \|N_j^{m-1}\|$ by Proposition 2.4. Hence, (c) holds.

Finally, we treat the case that λ_1 is the only eigenvalue of A. Let $U^*AU = \lambda_1 I_n + T$ be a Schur decomposition. The eigenprojection is $P_1 = I_n$ and the nilpotent operator is $N_1 = A-\lambda_1 I_n = UTU^*$. Since all eigenvalues of $A-\lambda_1 I_n$ are zero we have $\pi_\Lambda(A-\lambda_1 I_n) = 1$ by definition. If λ_1 is semisimple then also $\pi_\Sigma(A-\lambda_1 I_n) = \pi_\Sigma(0) = 1 = \pi_\Sigma(P_1)$. Hence, $q(A,\lambda_1) = 1$. Suppose $n \geq 2$ and λ_1 is nonderogatory. Then

$$q(A,\lambda_1) = \pi_\Sigma(A-\lambda_1 I_n) = \pi_\Sigma(T) = |\det(D)| = \|T^{n-1}\| = \|N_1^{n-1}\|,$$

where $T = \begin{bmatrix} \begin{smallmatrix}0\\ \vdots\end{smallmatrix} & D \\ \begin{smallmatrix}0 & \dots & 0\end{smallmatrix} & \end{bmatrix}$. □

4 Condition numbers

In the following $\mathcal{D}_\lambda(r)$ denotes the closed disk of radius $r > 0$ about $\lambda \in \mathbb{C}$. If $\lambda \in \Lambda(A)$, $A \in \mathbb{C}^{n\times n}$ then $\mathcal{C}_\lambda(\epsilon)$ denotes the connected component of the ϵ-pseudospectrum, $\Lambda_\epsilon(A)$, that contains λ. We define

$$\begin{aligned} R_\lambda^+(\epsilon) &:= \inf\{r>0 \mid \mathcal{C}_\lambda(\epsilon) \subseteq \mathcal{D}_\lambda(r)\ \}, \\ R_\lambda^-(\epsilon) &:= \sup\{r>0 \mid \mathcal{D}_\lambda(r)\ \subseteq \mathcal{C}_\lambda(\epsilon)\}, \end{aligned}$$

Then

$$\mathcal{D}_\lambda(R_\lambda^-(\epsilon)) \subseteq \mathcal{C}_\lambda(\epsilon) \subseteq \mathcal{D}_\lambda(R_\lambda^+(\epsilon)).$$

Theorem 4.1 *Let $\lambda \in \Lambda(A)$ be a nonderogatory eigenvalue of algebraic multiplicity m. Then*

$$R_\lambda^{\pm}(\epsilon) \;=\; q(A,\lambda)^{1/m}\,\epsilon^{1/m} + o(\epsilon^{1/m}). \tag{15}$$

The proof uses Theorem 1.1 and the lemma below.

Lemma 4.2 *Let $U \subseteq \mathbb{C}^n$ be an open neighborhood of $z_0 \in \mathbb{C}^n$. Let $f, g : U \to [0,\infty)$ be continuous functions. For $\epsilon \geq 0$ let $S_f(\epsilon)$ and $S_g(\epsilon)$ denote the connected component containing z_0 of the sublevel set $\{\, z \in U \mid f(z) \leq \epsilon \,\}$ and $\{\, z \in U \mid g(z) \leq \epsilon \,\}$ respectively. Assume that $0 = g(z_0)$ is an isolated zero of g, and*

$$\lim_{z\to z_0} \frac{f(z)}{g(z)} = 1. \tag{16}$$

Then there exists an $\epsilon_0 > 0$ and functions $h_\pm : [0,\epsilon_0] \to [0,\infty)$ with $\lim_{\epsilon\to 0} h_\pm(\epsilon) = 1$ such that for all $\epsilon \in [0,\epsilon_0]$,

$$S_g(\, h_-(\epsilon)\,\epsilon\,) \subseteq S_f(\epsilon) \subseteq S_g(\, h_+(\epsilon)\,\epsilon\,). \tag{17}$$

We postpone the proof of the lemma to the end of this section.

Proof of Theorem 4.1: Let in Lemma 4.2, $z_0 = \lambda$ and

$$f(z) = \sigma_{min}(A - zI_n), \qquad g(z) = \frac{|z-\lambda|^m}{q(A,\lambda)}, \qquad z \in \mathbb{C}.$$

Then $S_f(\epsilon) = \mathcal{C}_\lambda(\epsilon)$ and $S_g(\epsilon) = \mathcal{D}_\lambda((q(A,\lambda)\epsilon)^{1/m})$. Theorem 1.1 yields $\lim_{z\to\lambda} \frac{f(z)}{g(z)} = 1$. Hence, by the lemma there are functions $h_\pm$ with $\lim_{\epsilon\to 0} h_\pm(\epsilon) = 1$ and

$$\mathcal{D}_\lambda(\,(q(A,\lambda)\,h_-(\epsilon)\epsilon\,)^{1/m}\,) \subseteq \mathcal{C}_\lambda(\epsilon) \subseteq \mathcal{D}_\lambda(\,(q(A,\lambda)\,h_+(\epsilon)\epsilon\,)^{1/m}\,).$$

This shows (15). □

Now, we give the definition for the Hölder condition number of an eigenvalue of arbitrary multiplicity (see [4]). For $\lambda \in \mathbb{C}$, $m \in \mathbb{N}$ and $\widetilde{A} \in \mathbb{C}^{n\times n}$ we set

$$d_m(\widetilde{A},\lambda) := \min\{\, r \geq 0 \mid \mathcal{D}_\lambda(r) \text{ contains at least } m \text{ eigenvalues of } \widetilde{A}\,\}.$$

If λ is an eigenvalue of $A \in \mathbb{C}^{n\times n}$ of algebraic multiplicity m then the Hölder condition number of λ to the order $\alpha > 0$ is defined by

$$\mathrm{cond}_\alpha(A,\lambda) = \lim_{\epsilon\searrow 0} \sup_{\|\Delta\|\leq\epsilon} \frac{d_m(A+\Delta,\lambda)}{\|\Delta\|^\alpha}.$$

It is easily seen that $0 \neq \mathrm{cond}_\alpha(A,\lambda) \neq \infty$ for at most one order $\alpha > 0$.

Theorem 4.3 *Let $\lambda \in \Lambda(A)$ be a nonderogatory eigenvalue of multiplicity m. Then*

$$\mathrm{cond}_{1/m}(A,\lambda) \;=\; q(A,\lambda)^{1/m} = \begin{cases} \|P\| & \text{if } m=1, \\ \|N^{m-1}\|^{1/m} & \text{otherwise,} \end{cases} \tag{18}$$

where $P \in \mathbb{C}^{n\times n}$ is the eigenprojector onto the generalized eigenspace $\ker(A-\lambda I)^m$, *and* $N=(A-\lambda I)P$.

Proof: Let $\Delta \in \mathbb{C}^{n\times n}$ with $\|\Delta\| \le \epsilon$. Then the continuity of eigenvalues yields, that at least m eigenvalues of $A+\Delta$ are contained in $\mathcal{C}_\lambda(\epsilon)$. Hence

$$d_m(A+\Delta,\lambda) \;\le\; R^+_\lambda(\epsilon) = q(A,\lambda)^{1/m}\epsilon^{1/m} + o(\epsilon^{1/m}).$$

By letting $\epsilon = \|\Delta\|$ we obtain that for all $\Delta \in \mathbb{C}^{n\times n}$,

$$\frac{d_m(A+\Delta,\lambda)}{\|\Delta\|^{1/m}} \le q(A,\lambda)^{1/m} + o(\|\Delta\|^{1/m})\|\Delta\|^{-(1/m)}.$$

This yields

$$\mathrm{cond}_{1/m}(A,\lambda) \le q(A,\lambda)^{1/m}.$$

Let $z_\epsilon \in \mathbb{C}$ is a boundary point of $\mathcal{C}_\lambda(\epsilon)$. Then there is a Δ_ϵ with $\|\Delta_\epsilon\| = \epsilon$ and $z \in \Lambda(A+\Delta_\epsilon)$. If ϵ is such that $\mathcal{C}_\lambda(\epsilon) \cap \Lambda(A) = \{\lambda\}$ then precisely m eigenvalues of $A+\Delta_\epsilon$ are contained in $\mathcal{C}_\lambda(\epsilon)$. Thus,

$$\begin{aligned} d_m(A+\Delta_\epsilon,\lambda) &= |z_\epsilon - \lambda| \\ &\ge R^-_\lambda(\epsilon) \\ &= q(A,\lambda)^{1/m}\epsilon^{1/m} + o(\epsilon^{1/m}). \end{aligned}$$

and therefore

$$\frac{d_m(A+\Delta_\epsilon,\lambda)}{\|\Delta_\epsilon\|^{1/m}} \ge q(A,\lambda)^{1/m} + o(\epsilon^{1/m})\epsilon^{-(1/m)}.$$

Hence, $\mathrm{cond}_{1/m}(A,\lambda) \ge q(A,\lambda)^{1/m}$. □

Remark: In [6] (see also [4, 5]) the following generalization of Theorem 4.3 has been shown. Let λ be an *arbitrary* eigenvalue of A. If λ is semisimple then

$$\mathrm{cond}_1(A,\lambda) = \|P\|.$$

If λ is not semisimple then

$$\mathrm{cond}_{1/m}(A,\lambda) = \|N^{m-1}\|^{1/m},$$

where m denotes the index of nilpotency of N.

Proof of Lemma 4.2 By B_r we denote the closed ball of radius $r > 0$ about z_0. The condition that z_0 is an isolated zero of g combined with (16) yields that z_0 is also an isolated zero of f. Hence, there is an $r_0 > 0$ such that $f(z) > 0$ for all $z \in B_{r_0} \setminus \{z_0\}$. This implies that $\epsilon_r := \min_{z\in\partial B_r} f(z) > 0$ for any $r \in (0, r_0]$. If $\epsilon < \epsilon_r$ then ∂B_r does not intersect the sublevel sets $\{\, z \in U \mid \ f(z) \leq \epsilon \,\}$. Thus $S_f(\epsilon)$ is contained in the interior of B_r. Note that $S_f(\epsilon)$ being a connected component of a closed set is closed. It follows that $S_f(\epsilon)$ is compact if $\epsilon < \epsilon_{r_0}$. Now, let

$$\phi_\pm(z) := \begin{cases} (1 \pm \|z - z_0\|)\frac{g(z)}{f(z)} & z \in B_{r_0} \setminus \{z_0\}, \\ 1, & z = z_0. \end{cases}$$

Condition (16) yields that the functions $\phi_\pm : U \to \mathbb{R}$ are continuous. For $\epsilon < \epsilon_{r_0}$ let

$$h_-(\epsilon) := \min_{z\in S_f(\epsilon)} \phi_-(z), \quad h_+(\epsilon) := \max_{z\in S_f(\epsilon)} \phi_+(z).$$

Then we have for all $\epsilon < \epsilon_r$,

$$\min_{z\in\partial B_r} \phi_\pm(z) \leq h_\pm(\epsilon) \leq \max_{z\in\partial B_r} \phi_\pm(z).$$

As r tends to 0 the max and the min tend to $\phi_\pm(z_0) = 1$. This yields $\lim_{\epsilon\to 0} h_\pm(\epsilon) = 1$. If $z \in \partial S_f(\epsilon)$ then $f(z) = \epsilon$ and $g(z) > (1 - \|z - z_0\|)\frac{g(z)}{f(z)} f(z) \geq h_-(\epsilon)\epsilon$. Thus $\partial S_f(\epsilon)$ does not intersect $E := \{\, z \in U \mid \ g(z) \leq h_-(\epsilon)\epsilon \,\}$. Thus $S_g(h_-(\epsilon)\epsilon)$ being a connected component of E is either contained in the interior of $S_f(\epsilon)$ or in the complement of $S_f(\epsilon)$. The latter is impossible since $z_0 \in S_f(\epsilon) \cap S_g(h_-(\epsilon)\epsilon)$. Hence, $S_g(h_-(\epsilon)\epsilon) \subset S_f(\epsilon)$. This proves the first inclusion in (17). To prove the second suppose $z_0 \neq z \in \partial S_g(h_+(\epsilon)\epsilon) \cap S_f(\epsilon)$. Then $g(z) = h_+(\epsilon)\epsilon$ and $0 < f(z) \leq \epsilon$. Hence $g(z)/f(z) \geq h_+(\epsilon)$, a contradiction. Thus $S_f(\epsilon)$ is contained in the interior of $S_g(h_+(\epsilon)\epsilon)$. □

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] Hinrichsen, D.; Pritchard, A.J.: Dynamical Systems Theory. Manuscript, 2004.

[3] Horn, R.A.; Johnson, C.R.: Matrix analysis. Cambridge University Press, 1985.

[4] Chaitin-Chatelin, F.; Harrabi, A.; Ilahi, A.: About Hölder condition numbers and the stratification diagram for defective eigenvalues. Math. Comput. Simul. 54, No.4-5, 397-402 (2000)

[5] Harrabi, A.: *Pseudospectres d'Operateurs Intégraux et Différentiels: Application a la Physique Mathematique.* Thesis. Universite des Sciences Sociales de Toulouse. May 1998.

[6] Karow, M.: Geometry of spectral value sets. Ph.D. thesis. Universität Bremen, July 2003.

[7] Moro, J.; Burke, J.V.; Overton, M. L.: On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. *SIAM J. Matrix Anal. Appl.* 18(4):793-817, 1997.

[8] Stewart, G.W., Sun, J.: MATRIX PERTURBATION THEORY. Academic Press, San Diego, 1990.