Second order pseudospectra of normal matrices*

Gorka Armentia[†], Juan-Miguel Gracia[‡], Francisco E. Velasco[‡]

November 28, 2010

Abstract

Let $A \in \mathbb{C}^{n\times n}$ be a normal matrix. A well-known theorem asserts that for any $\varepsilon \geq 0$ the ordinary ε-pseudospectrum of A is the union of the closed disks of radius ε centered at the eigenvalues of A. We give a proof of the converse theorem.

We define the ε-pseudospectrum of algebraic (resp. geometric) multiplicity ≥ 2 of A as the set of complex numbers z such that there exists a $X \in \mathbb{C}^{n\times n}$ which satisfies $\|X - A\| \leq \varepsilon$ and z is an eigenvalue of X of algebraic (resp. geometric) multiplicity ≥ 2. Let us denote this set by $\Lambda^{(\mathrm{a})}_{\varepsilon,2}(A)$, (resp. $\Lambda^{(\mathrm{g})}_{\varepsilon,2}(A)$). Here, $\|\cdot\|$ stands for the spectral norm.

We prove that $\Lambda^{(\mathrm{a})}_{\varepsilon,2}(A)$ is a union of closed disks, whose centers and radii are determined in terms of the eigenvalues of A and ε. Also, we prove that $\Lambda^{(\mathrm{g})}_{\varepsilon,2}(A)$ is the union of pairwise intersections of closed disks centered at the eigenvalues of A and radius ε.

AMS classification: 15A18, 15A21, 15A60, 47A25.

Key Words: singular values, unitary similarity, algebraic multiplicity, geometric multiplicity, Malyshev.

1 Introduction

For any complex matrix M we denote by $\sigma_1(M) \geq \sigma_2(M) \geq \cdots$ its singular values decreasingly ordered. We write $\|\cdot\|$ for the spectral matrix norm. Given a matrix $A \in \mathbb{C}^{n\times n}$, we denote by $\Lambda(A)$ the spectrum or set of eigenvalues of A. Let us recall the concept of unordered n-tuple of complex numbers. We say that two n-tuples $(\lambda_1, \lambda_2, \dots, \lambda_n)$, $(\mu_1, \mu_2, \dots, \mu_n) \in \mathbb{C}^n$ are equivalent if there exists a permutation π of the set $\{1, 2, \dots, n\}$ such that

$$(\mu_1, \mu_2, \dots, \mu_n) = (\lambda_{\pi(1)}, \lambda_{\pi(2)}, \dots, \lambda_{\pi(n)}).$$

This defines a equivalence relation $\sim$ in $\mathbb{C}^n$. We denote by $\mathbb{C}^n_{\mathrm{sym}}$ the quotient set $\mathbb{C}^n/\sim$. We call unordered n-tuple of complex numbers to each element $[(\lambda_1, \lambda_2, \dots, \lambda_n)] \in \mathbb{C}^n_{\mathrm{sym}}$, which we write briefly by $(\lambda_1, \lambda_2, \dots, \lambda_n)$ without

*This paper has been written financied by the MICINN Spanish Ministry MTM2010-19356-C02-01, and the Basque Government GIC10/169-IT361-10.

[†]Department of Mathematical Engineering and Computer Science, The Public University of Navarre, Campus de Arrosadía, 31006 Pamplona, Spain, gorka.armentia@unavarra.es

[‡]Department of Applied Mathematics and Statistics and R.O., The University of the Basque Country, Faculty of Pharmacy, Paseo de la Universidad, 7; 01006 Vitoria-Gasteiz, Spain, juanmiguel.gracia@ehu.es, franciscoenrique.velasco@ehu.es

peril of confusion. Now, if $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of a matrix $A \in \mathbb{C}^{n\times n}$ where the multiples appear repeated, we denote by $\mathrm{Eig}(A)$ the unordered $n-$tuple $(\lambda_1, \lambda_2, \dots, \lambda_n)$. Thus, $\mathrm{Eig}(A) \in \mathbb{C}^n_{\mathrm{sym}}$. We will use $\Lambda(A)$ and $\mathrm{Eig}(A)$. They are different things. The spectrum $\Lambda(A)$ is a subset of $\mathbb{C}$ and may have less than n elements. See [3, p. 153]

For $\varepsilon \geq 0$, it is defined the ordinary $\varepsilon-$*pseudospectrum* of A as

$$\Lambda_\varepsilon(A) := \bigcup_{\|X-A\|\leq\varepsilon} \Lambda(X).$$

Let us note that $\Lambda_0(A) = \Lambda(A)$. The set of eigenvalues of a matrix A with algebraic multiplicity greater than or equal to k will be denoted by $\Lambda_k^{(\mathrm{a})}(A)$. So, for $\varepsilon \geq 0$, we define the $\varepsilon-$*pseudospectrum* of algebraic multiplicity $\geq k$ of A as the set

$$\Lambda_{\varepsilon,k}^{(\mathrm{a})}(A) := \bigcup_{\|X-A\|\leq\varepsilon} \Lambda_k^{(\mathrm{a})}(X).$$

Also, briefly we will call this pseudospectrum, the *algebraic* ε*-pseudospectrum of order* k of A. Let us observe again that $\Lambda_{0,k}^{(\mathrm{a})}(A) = \Lambda_k^{(\mathrm{a})}(A)$.

It is well known that the ε-pseudospectrum of A can be characterized as the sublevel set ε of the function $z \mapsto \sigma_n(zI_n - A)$, which associates to each complex z the nth singular value of the matrix $zI_n - A$. In particular, we have $\Lambda_\varepsilon(A) = \{z \in \mathbb{C} \colon \sigma_n(zI_n - A) \leq \varepsilon\}$.

Other analogous characterization of the set $\Lambda_{\varepsilon,2}^{(\mathrm{a})}(A)$ can be deduced from the next result due to Malyshev [8].

Theorem 1. *Given $A \in \mathbb{C}^{n\times n}$ and $z \in \mathbb{C}$, let us denote by*

$$h_2(z) := \max_{t\geq 0} \sigma_{2n-1} \begin{pmatrix} zI_n - A & tI_n \\ 0 & zI_n - A \end{pmatrix}. \tag{1}$$

Then, we have

$$\min_{z\in\Lambda_2^{(\mathrm{a})}(X)} \|X - A\| = h_2(z).$$

With this result, we get the following characterization of $\varepsilon-$pseudospectrum of A with algebraic multiplicity ≥ 2.

Proposition 2. *Given $A \in \mathbb{C}^{n\times n}$, for all $\varepsilon \geq 0$ $z \in \mathbb{C}$ we have*

$$\Lambda_{\varepsilon,2}^{(\mathrm{a})}(A) = \{z \in \mathbb{C} : h_2(z) \leq \varepsilon\}.$$

Proof. Let be $z \in \Lambda_{\varepsilon,2}^{(\mathrm{a})}(A)$. Then, there exists a matrix $Y \in \mathbb{C}^{n\times n}$ such that $z \in \Lambda_2^{(\mathrm{a})}(Y)$ and, in addition, $\|Y - A\| \leq \varepsilon$. Therefore, following on from (1), we derive that

$$h_2(z) = \min_{z\in\Lambda_2^{(\mathrm{a})}(X)} \|X - A\| \leq \|Y - A\| \leq \varepsilon.$$

Reciprocally, let z be such that $h_2(z) \leq \varepsilon$. Then by Theorem 1 we have $\min_{z\in\Lambda_2^{(\mathrm{a})}(X)} \|X - A\| \leq \varepsilon$. Therefore, there exists a matrix Y such that $z \in \Lambda_2^{(\mathrm{a})}(Y)$ and, in addition, $\|Y - A\| \leq \varepsilon$. So, $z \in \Lambda_{\varepsilon,2}^{(\mathrm{a})}(A)$. □

The goal of this paper is to describe the set $\Lambda_{\varepsilon,2}^{(\mathrm{a})}(A)$, as long as the matrix $A \in \mathbb{C}^{n\times n}$ is normal. So, denoting by

$$\mathcal{D}(\lambda, \delta) := \{z \in \mathbb{C} :\mid z - \lambda \mid \leq \delta\},$$

The main result of this paper is the following theorem.

Theorem 3. *Let us suppose that the matrix $A \in \mathbb{C}^{n\times n}$ is normal. Moreover, let us suppose that* $\mathrm{Eig}(A) = (\lambda_1, \lambda_2, \ldots, \lambda_n)$. *Let $\varepsilon \geq 0$ and (j, k) a pair of integers such that $1 \leq j < k \leq n$. If $\varepsilon \geq |\lambda_j - \lambda_k|/2$, let us define*

$$\rho_{jk}(\varepsilon) := \sqrt{\varepsilon^2 - \frac{|\lambda_j - \lambda_k|^2}{4}}, \tag{2}$$

$$\mathcal{D}_{jk}(\varepsilon) := \mathcal{D}\left(\frac{\lambda_j + \lambda_k}{2}, \rho_{jk}(\varepsilon)\right). \tag{3}$$

If $\varepsilon < |\lambda_j - \lambda_k|/2$, let us define $\mathcal{D}_{jk}(\varepsilon) := \emptyset$.
Then

$$\Lambda_{\varepsilon,2}^{(\mathrm{a})}(A) = \bigcup_{1\leq j<k\leq n} \mathcal{D}_{jk}(\varepsilon).$$

Remark 1. Since A is normal, then there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $U^*AU = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Moreover, since U is unitary, we have $\Lambda_{\varepsilon,2}^{(\mathrm{a})}(A) = \Lambda_{\varepsilon,2}^{(\mathrm{a})}(U^*AU)$. So, from now on, we will assume without loss of generality $A = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.

The organization of this paper is as follows: in Section 2, we revise a characterization of the normal matrices $A \in \mathbb{C}^{n\times n}$ as those for which the set $\Lambda_\varepsilon(A)$ is the union of closed disks of radius ε with center at its eigenvalues, for every $\varepsilon > 0$. In Section 3, we introduce the necessary results to show the theorem 3. We prove the theorem 3 in Section 4. Finally, in Section 5, we describe the ε-pseudospectrum of A of geometric multiplicity $\geq k$, when the matrix A is normal.

2 Matrices whose ordinary pseudospectra are union of disks

In this section, we revise a result that characterizes the normal matrices as those whose ordinary ε-pseudospectrum, for each $\varepsilon \geq 0$, is the union of closed disks of radius ε centered at the eigenvalues. Let $A \in \mathbb{C}^{n\times n}$ be a normal matrix, that is, $A^*A = AA^*$, where A^* stands for the conjugate transpose of A. The following theorem appears in [9, Theorem 2.2, p. 19] and [7, Corollary 5.3.8, p. 73].

Theorem 4. *Let $A \in \mathbb{C}^{n\times n}$, $\Lambda(A) = \{\lambda_1, \ldots, \lambda_v\}$. Then the matrix A is normal if and only if, for each $\varepsilon > 0$,*

$$\Lambda_\varepsilon(A) = \bigcup_{j=1}^{v} \mathcal{D}(\lambda_j, \varepsilon). \tag{4}$$

As the proof was given in [9, Theorem 2.2, p. 19] only for the case in which the eigenvalues of A are simple, and [7, Corollary 5.3.8, p. 73] proved the "only if" part, we give a proof for reasons of completeness.

We need some previous lemmas. The first is a characterization of a normal matrix by its Jordan decomposition. Let

$$A = \sum_{j=1}^{v} (\lambda_j P_j + N_j)$$

be the Jordan decomposition of any $A \in \mathbb{C}^{n\times n}$ whose spectrum is $\{\lambda_1, \ldots, \lambda_v\}$, where, for each $j \in \{1, \ldots, v\}$, P_j is the Riesz projector onto the root subspace (or generalized eigenspace) $\mathcal{R}_{\lambda_j}(A)$ of λ_j and along the sum of root subspaces associated with all eigenvalues of A different from λ_j; and $N_j := (A - \lambda_j I_n)P_j$ is the nilpotent matrix corresponding to λ_j. So, $P_i P_j = \delta_{ij} P_j$ and $\sum_{j=1}^{v} P_j = I_n$.

Lemma 5. *Let $A \in \mathbb{C}^{n\times n}$ and $\Lambda(A) = \{\lambda_1, \ldots, \lambda_v\}$. Then, A is normal if and only if*

$$A = \sum_{j=1}^{v} \lambda_j P_j,$$

with $P_i P_j = \delta_{ij} P_j$ and $P_j^* = P_j$.

A proof of this lemma can be seen in [10, Theorem 8.1 (10), p. 241].

Let us fix $j \in \{1, \ldots, v\}$ and let us denote by $\mathcal{K}_j(\varepsilon)$ the connected component of $\Lambda_\varepsilon(A)$ that contains λ_j for each $\varepsilon \geq 0$. We denote by ν_j and m_j the index and the algebraic multiplicity of λ_j, respectively. Let us recall that the spectral variation of $A' \in \mathbb{C}^{n\times n}$ with respect to the eigenvalue λ_j of A, denoted by $\mathrm{sv}_{(A,\lambda_j)}(A')$, is defined as the minimum of $\rho \geq 0$ such that the closed disk $\mathcal{D}(\lambda_j, \rho)$ with center at λ_j and radius ρ contains at least m_j eigenvalues (counting multiplicities) of A'. Moreover, the (Hölder) condition number of the eigenvalue λ_j of order $1/\nu_j$ is defined by

$$c_{1/\nu_j}(\lambda_j) := \lim_{\varepsilon \to 0^+} \max_{0 < \|X - A\| \leq \varepsilon} \frac{\mathrm{sv}_{(A,\lambda_j)}(X)}{\|X - A\|^{1/\nu_j}}.$$

By [7, Theorem 5.4.4 (iv),(viii), p. 77–78], we have the following lemma.

Lemma 6.

$$c_{1/\nu_j}(\lambda_j) = \begin{cases} \|P_j\|, & \text{if } \nu_j = 1, \\ \|N_j^{\nu_j - 1}\|^{1/\nu_j} & \text{if } \nu_j \geq 2. \end{cases}$$

Lemma 7. *Let $\delta_j(\varepsilon)$ be the diameter of $\mathcal{K}_j(\varepsilon)$ for each $\varepsilon \geq 0$. Then*

$$\delta'(0^+) = \begin{cases} 2\, c_1(\lambda_j), & \text{if } \nu_j = 1, \\ \infty, & \text{if } \nu_j \geq 2. \end{cases}$$

The proof of this Lemma can be seen in [2, Theorem 4.2]. By O or O_t we denote a zero matrix or the $t \times t$ zero matrix. Let us recall that a projector is a matrix $P \in \mathbb{C}^{n\times n}$ such that $P^2 = P$. The next lemma uses a results on the singular values of a projector.

Lemma 8. *Let $P \in \mathbb{C}^{n\times n}, P \neq O$, a projector such that $\|P\| = 1$. Then $P^* = P$.*

Proof. From [5, p. 235] we deduce that if σ is a singular value of P, then either $\sigma = 0$ or else $\sigma \geq 1$. See also [1]. By a result of Djoković [4, Theorem 4] there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that either

$$U^*PU = \begin{pmatrix} 1 & s_1 \\ 0 & 0 \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} 1 & s_k \\ 0 & 0 \end{pmatrix} \oplus I_m \oplus O_t,$$

if P has singular values greater than 1, or

$$U^*PU = I_m \oplus O_t,$$

otherwise.

Therefore, if $\|P\| = 1$, then the singular values of P are 1 and 0. So, $P = U(I_m \oplus O_t)U^*$. Consequently, $P^* = P$.

□

Proof of Theorem 4. If A is a normal matrix, then there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $U^*AU = \mathrm{diag}(\mu_1, \ldots, \mu_n)$. As $\Lambda_\varepsilon(A) = \Lambda_\varepsilon(U^*AU)$, we deduce that

$$\Lambda_\varepsilon(A) = \Lambda_\varepsilon\Big(\mathrm{diag}(\mu_1, \ldots, \mu_n)\Big) = \bigcup_{j=1}^{n} \mathcal{D}(\mu_j, \varepsilon).$$

Conversely, if for each $\varepsilon > 0$

$$\Lambda_\varepsilon(A) = \bigcup_{j=1}^{v} \mathcal{D}(\lambda_j, \varepsilon),$$

then for each eigenvalue λ_j of A and each sufficiently small $\varepsilon \geq 0$ we have $\mathcal{K}_j(\varepsilon) = \mathcal{D}(\lambda_j, \varepsilon)$. Thus, the diameter $\delta_j(\varepsilon)$ of $\mathcal{K}_j(\varepsilon)$ is equal to $2\,\varepsilon$. Therefore, $\delta_j'(\varepsilon) = 2$, and $\delta_j'(0^+) = 2$. By Lemma 7, this implies that the index ν_j of λ_j is equal to 1. So, $c_1(\lambda_j) = 1$, which implies that $\|P_j\| = 1$. From Lemma 8, $P_j^* = P_j$. Therefore, the matrix A is normal.

□

3 Necessary results to prove Theorem 3

Given $z \in \mathbb{C}$, in this section we consider the matrix function

$$F_z(t) := \begin{pmatrix} zI_n - A & tI_n \\ 0 & zI_n - A \end{pmatrix},$$

defined for every $t \geq 0$, with $A = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. For that function, the singular values are given by the following result.

Proposition 9. *Given $z \in \mathbb{C}$, the singular values of $F_z(t)$ not necessarily ordered are*

$$f_{z,1}^-(t), f_{z,1}^+(t), f_{z,2}^-(t), f_{z,2}^+(t), \ldots, f_{z,n}^-(t), f_{z,n}^+(t),$$

where, for each $i = 1, 2, \ldots, n$

$$\begin{cases} f^-_{z,i}(t) = \sqrt{\dfrac{t^2 + 2|z-\lambda_i|^2 - \sqrt{t^4 + 4t^2|z-\lambda_i|^2}}{2}} \\ f^+_{z,i}(t) = \sqrt{\dfrac{t^2 + 2|z-\lambda_i|^2 + \sqrt{t^4 + 4t^2|z-\lambda_i|^2}}{2}}. \end{cases} \tag{5}$$

Proof. It is enough to observe that, permuting rows and columns properly, the system of singular values of $F_z(t)$ fits in with that of the matrix

$$\bigoplus_{i=1}^{n} \begin{pmatrix} z-\lambda_i & t \\ 0 & z-\lambda_i \end{pmatrix}.$$

□

Now, let us see some properties of the functions $f^-_{z,i}(t)$ and $f^+_{z,i}(t)$, defined for every $t \geq 0$.

Proposition 10. *For $i = 1, 2, \ldots, n$, we have:*
(a) $f^-_{z,i}(0) = f^+_{z,i}(0) = |z - \lambda_i|$.
(b) *If $z = \lambda_j$, then $f^-_{z,j}(t) = 0$ for $t \geq 0$.*
(c)

$$\lim_{t\to\infty} f^+_{z,i}(t) = \infty, \text{ and } \lim_{t\to\infty} f^-_{z,i}(t) = 0.$$

(d) *$f^+_{z,i}(t)$ is increasing.*
(e) *$f^-_{z,i}(t)$ is decreasing.*

Proof. The statements (a), (b), (c) and (d) are immediate by (5). So, it remains to prove that the function $f^-_{z,i}(t)$ is decreasing or, equivalently, that the function

$$2f^-_{z,i}(t)^2 = t^2 + 2|z-\lambda_i|^2 - \sqrt{t^4 + 4t^2|z-\lambda_i|^2}$$

is decreasing, or what is the same, that the function

$$g(t) := t^2 - \sqrt{t^4 + 4t^2a^2},$$

is decreasing for $t \geq 0$. By differentiating, for $t > 0$ we have

$$g'(t) \leq 0 \Leftrightarrow 2t \leq \frac{2t^3 + 4ta^2}{\sqrt{t^4 + 4t^2a^2}} \Leftrightarrow 1 \leq \frac{t^2 + 2a^2}{\sqrt{t^4 + 4t^2a^2}} \Leftrightarrow t^4 + 4t^2a^2 \leq (t^2 + 2a^2)^2,$$

a fact which is true.

□

Proposition 11. *If $|z - \lambda_j| \leq |z - \lambda_k|$, then for $t \geq 0$ we have:*
(a) $f^+_{z,j}(t) \leq f^+_{z,k}(t)$.
(b) $f^-_{z,j}(t) \leq f^-_{z,k}(t)$.

Proof. (a) is immediately deduced from (5). To prove (b), since $f^-_{z,j}(t) \leq f^-_{z,k}(t) \Leftrightarrow 2f^-_{z,j}(t)^2 \leq 2f^-_{z,k}(t)^2$, following from (5), it is enough to prove that the function

$$g(x) := 2x^2 - \sqrt{t^4 + 4t^2x^2}$$

is increasing for $x \geq 0$. By differentiating, for $x > 0$ we find that

$$g'(x) \geq 0 \Leftrightarrow 4x \geq \frac{4t^2x}{\sqrt{t^4+4t^2x^2}} \Leftrightarrow 1 \geq \frac{t^2}{\sqrt{t^4+4t^2x^2}},$$

which is true.

□

Proposition 12. *Let us suppose that* $|z-\lambda_j| \leq |z-\lambda_k|$. *Then* $f^+_{z,j}(t) = f^-_{z,k}(t)$ *if and only if* $t = t^z_{jk}$, *where*

$$t^z_{jk} = \begin{cases} 0 & \text{if } \lambda_j = \lambda_k \\ \dfrac{|z-\lambda_k|^2 - |z-\lambda_j|^2}{\sqrt{2\left(|z-\lambda_k|^2 + |z-\lambda_j|^2\right)}} & \text{if } \lambda_j \neq \lambda_k \end{cases} \tag{6}$$

Proof. The case $|z-\lambda_j| = |z-\lambda_k|$ is immediately deduced from (5). Now, let us suppose that $|z-\lambda_j| < |z-\lambda_k|$. Since $f^+_{z,j}(t) = f^-_{z,k}(t)$, if and only if $2f^+_{z,j}(t)^2 = 2f^+_{z,k}(t)^2$, following from (5), we conclude that $f^+_{z,j}(t) = f^-_{z,k}(t)$ if and only if

$$2|z-\lambda_j|^2 + \sqrt{t^4+4t^2|z-\lambda_j|^2} = 2|z-\lambda_k|^2 - \sqrt{t^4+4t^2|z-\lambda_k|^2} \Leftrightarrow$$
$$\sqrt{t^4+4t^2|z-\lambda_j|^2} = 2(|z-\lambda_k|^2 - |z-\lambda_j|^2) - \sqrt{t^4+4t^2|z-\lambda_k|^2}.$$

Now, denoting by $\alpha := 2(|z-\lambda_k|^2 - |z-\lambda_j|^2)$, which is positive, and squaring we infer that $f^+_{z,j}(t) = f^-_{z,k}(t)$ if and only if

$$t^4 + 4t^2|z-\lambda_j|^2 = \alpha^2 + t^4 + 4t^2|z-\lambda_k|^2 - 2\alpha\sqrt{t^4+4t^2|z-\lambda_k|^2} \Leftrightarrow$$
$$2\alpha\sqrt{t^4+4t^2|z-\lambda_k|^2} = \alpha^2 + 4t^2(|z-\lambda_k|^2 - |z-\lambda_j|^2) = \alpha^2 + 2\alpha t^2 \Leftrightarrow$$
$$2\sqrt{t^4+4t^2|z-\lambda_k|^2} = \alpha + 2t^2 \Leftrightarrow 4t^4 + 16t^2|z-\lambda_k|^2 = \alpha^2 + 4t^4 + 4\alpha t^2.$$

That is,

$$t^2 = \frac{\alpha^2}{16|z-\lambda_k|^2 - 4\alpha} = \frac{\left(|z-\lambda_k|^2 - |z-\lambda_j|^2\right)^2}{2\left(|z-\lambda_k|^2 + |z-\lambda_j|^2\right)}.$$

□

The next proposition describes the singular value $\sigma_{2n-1}(F_z(t))$.

Proposition 13. *Let us suppose*

$$|z-\lambda_j| \leq |z-\lambda_k| \leq |z-\lambda_i|,\ i = 1,2,\ldots,n,\ i \neq j,k.$$

Then,

$$\sigma_{2n-1}(F_z(t)) = \min\{f^+_{z,j}(t), f^-_{z,k}(t)\}.$$

Moreover, if $\lambda_j = \lambda_k$ *then,*

$$\sigma_{2n-1}(F_z(t)) = f^-_{z,j}(t).$$

Proof. First, let us observe that, since $|z-\lambda_j| \leq |z-\lambda_i|$ for $i \neq j$, by Proposition 11, we have for $t \geq 0$

$$f_{z,j}^-(t) \leq f_{z,i}^-(t), \text{ and } f_{z,j}^-(t) \leq f_{z,j}^+(t) \leq f_{z,i}^+(t),$$

for every $i \neq j$. So,

$$\sigma_{2n}(F_z(t)) = f_{z,j}^-(t). \tag{7}$$

Now, let t_{jk}^z be the point where $f_{z,j}^+(t) = f_{z,k}^-(t)$. Then, since $f_{z,j}^+(t)$ is increasing and $f_{z,k}^-(t)$ is decreasing (Proposition 10), and furthermore, $f_{z,j}^+(t_{jk}^z) = f_{z,k}^-(t_{jk}^z)$, we conclude that $f_{z,j}^+(t) \leq f_{z,k}^-(t)$ in the interval $[0, t_{jk}^z]$. Now, following from (7) and Proposition 11, we find that

$$\sigma_{2n}(F_z(t)) = f_{z,j}^-(t) \leq f_{z,j}^+(t) \leq f_{z,k}^-(t) \leq \begin{cases} f_{z,i}^-(t) & i \neq j,k \\ f_{z,i}^+(t) & i \neq j. \end{cases}$$

Consequently, $\sigma_{2n-1}(F_z(t)) = f_{z,j}^+(t)$, for $t \in [0, t_{jk}^z]$. Now, let us suppose that $t \in [t_{jk}^z, \infty)$. Then, since $f_{z,j}^+(t)$ is increasing and $f_{z,k}^-(t)$ is decreasing (Proposition 10), and, moreover $f_{z,j}^+(t_{jk}^z) = f_{z,k}^-(t_{jk}^z)$, we infer that $f_{z,j}^+(t) \geq f_{z,k}^-(t)$ in the interval $[t_{jk}^z, \infty)$. Now, following from (7) and Proposition 11, we have

$$\sigma_{2n}(F_z(t)) = f_{z,j}^-(t) \leq f_{z,k}^-(t) \leq \begin{cases} f_{z,j}^+(t) \leq f_{z,i}^+(t) & i \neq j,k \\ f_{z,i}^-(t) & i \neq j,k. \end{cases}$$

Therefore, $\sigma_{2n-1}(F_z(t)) = f_{z,k}^-(t)$, for $t \in [t_{jk}^z, \infty)$. To conclude, let us observe that if $\lambda_j = \lambda_k$, then $t_{jk}^z = 0$ and $\sigma_{2n-1}(F_z(t)) = f_{z,k}^-(t) = f_{z,j}^-(t)$.

□

For the following property we introduce some notations. Given $z \in \mathbb{C}$ and λ_j, λ_k two eigenvalues of A, let us suppose that $|z-\lambda_j| \leq |z-\lambda_k|$. With these considerations, we define

$$g_{jk}(z,t) := \min\{f_{z,j}^+(t), f_{z,k}^-(t)\}, \tag{8}$$

$$g_{jk}^z := \max_{t \geq 0} g_{jk}(z,t). \tag{9}$$

With these notations, we find the following result.

Proposition 14. *If t_{jk}^z is defined in* (6)*, then $g_{jk}^z = g_{jk}(z, t_{jk}^z)$.*

Proof. First, let us observe that, by Proposition 12, we have $f_{z,j}^+(t_{jk}^z) = f_{z,k}^-(t_{jk}^z)$. Now, as in the proof of Proposition 13, we deduce that $f_{z,j}^+(t) \leq f_{z,k}^-(t)$ in the interval $[0, t_{jk}^z]$, and $f_{z,k}^-(t) \leq f_{z,j}^+(t)$ in $[t_{jk}^z, \infty)$. So, the proof concludes by using the fact that the function $f_{z,j}^+(t)$ is increasing and that $f_{z,k}^-(t)$ is decreasing (Proposition 10).

□

Proposition 15. *Let us assume that $|z-\lambda_j| \leq |z-\lambda_k| \leq |z-\lambda_m|$. Then, we have:*
(a) *For every $t \geq 0$, $g_{jk}(z,t) \leq g_{km}(z,t)$ and $g_{jk}(z,t) \leq g_{jm}(z,t)$.*
(b) $g_{jk}^z \leq g_{km}^z$ *and* $g_{jk}^z \leq g_{jm}^z$.

Proof. Since $|z-\lambda_j| \leq |z-\lambda_k| \leq |z-\lambda_m|$, from Proposition 11 we infer that for $t \geq 0$,

$$f_{z,j}^{-}(t) \leq f_{z,k}^{-}(t) \leq f_{z,m}^{-}(t), \text{ and } f_{z,j}^{+}(t) \leq f_{z,k}^{+}(t) \leq f_{z,m}^{+}(t).$$

So

$$\min\{f_{z,j}^{+}(t), f_{z,k}^{-}(t)\} \leq \min\{f_{z,k}^{+}(t), f_{z,m}^{-}(t)\},$$

and

$$\min\{f_{z,j}^{+}(t), f_{z,k}^{-}(t)\} \leq \min\{f_{z,j}^{+}(t), f_{z,m}^{-}(t)\}.$$

Consequently, (a) follows from the expression given for $g_{pq}(z,t)$ in (8). (b) is immediate by (a).

□

To conclude this section, let us see a relation between $h_2(z)$ and g_{jk}^z, which are defined in (1) and (9), respectively.

Proposition 16. *Given $z \in \mathbb{C}$ and λ_j, λ_k two eigenvalues of A, let us suppose that $|z-\lambda_j| \leq |z-\lambda_k|$. Then,*
(a) *If $|z-\lambda_k| \leq |z-\lambda_i|$, $i \neq j,k$, then $h_2(z) = g_{jk}^z$.*
(b) *$h_2(z) \leq g_{jk}^z$.*

Proof. First, let us observe that if λ_p, λ_q are eigenvalues of A such that

$$|z-\lambda_p| \leq |z-\lambda_q| \leq |z-\lambda_i|, \; i \neq p,q,$$

then, from Proposition 13 and (8), we have

$$\sigma_{2n-1}\left(F_z(t)\right) = \min\{f_{z,p}^{+}(t), f_{z,q}^{-}(t)\} = g_{pq}(z,t).$$

Consequently, $h_2(z) = g_{pq}^z$. This proves the statement (a).

To prove (b), let us consider some situations: (1) $j=p, k=q$; (2) $j=p, k \neq q$; (3) $j=q$; (4)$j \neq p,q$. The case (1) is immediate by the statement (a). For (2), since $|z-\lambda_j| \leq |z-\lambda_q| \leq |z-\lambda_k|$, then by Proposition 15 and the statement (a), we infer that $h_2(z) = g_{jq}^z \leq g_{jk}^z$. The cases (3) and (4) are proved in a similar way.

□

4 Proof of Theorem 3

In the fourth section, we proceed with the proof of Theorem 3. In order to achieve it, the next result is key.

Theorem 17. *Let λ_j, λ_k be terms of* $\mathrm{Eig}(A)$. *Let us assume that $\varepsilon \geq \frac{|\lambda_k-\lambda_j|}{2}$. Let us define the midpoint of the segment of vertices λ_j and λ_k,*

$$z_0 := \frac{\lambda_j+\lambda_k}{2}.$$

Then, for each $z \in \mathbb{C}$ such that $|z-z_0| = \rho_{jk}(\varepsilon)$, where $\rho_{jk}(\varepsilon)$ is defined in (2), if $|z-\lambda_j| \leq |z-\lambda_k|$ or $|z-\lambda_k| \leq |z-\lambda_j|$ then

$$g_{jk}^z = \varepsilon \text{ or } g_{kj}^z = \varepsilon, \text{respectively.} \tag{10}$$

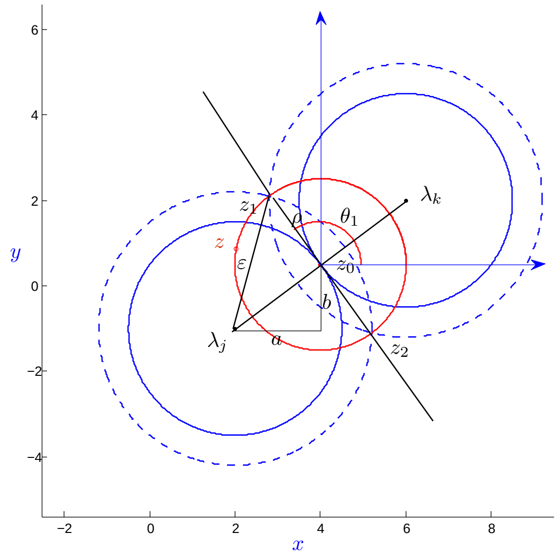

Figure 1: In a neighborhood of the midpoint z_0 between two close eigenvalues λ_j and λ_k. Moreover, $z_1 = z_0 + \rho e^{i\theta_1}$ and $z_2 = z_0 + \rho e^{i(\theta_1+\pi)}$ are the points of intersection of the circles centered at λ_j and λ_k and radius ε.

Proof. To begin with, we assume that $\lambda_j \neq \lambda_k$. Otherwise, the proof is easy. To shorten notation, we write ρ instead of $\rho_{jk}(\varepsilon)$ and

$$\begin{cases} a := \operatorname{Re} \frac{\lambda_k - \lambda_j}{2}, \ b := \operatorname{Im} \frac{\lambda_k - \lambda_j}{2}, \\ \theta_1 := \pi/2 + \arctan \frac{b}{a}, \ \text{if } a \neq 0; \quad \theta_1 := \pi, \ \text{if } a = 0, \\ \alpha(\theta) := a\cos\theta + b\sin\theta, \ \theta_1 \leq \theta \leq \theta_1 + \pi. \end{cases} \tag{11}$$

It is easy to demonstrate that for each $\theta \in (\theta_1, \theta_1 + \pi)$, $\alpha(\theta) < 0$; also, $\alpha(\theta_1) = 0 = \alpha(\theta_1 + \pi)$. Let us note that, following from expression (2) we infer that

$$\varepsilon^2 := a^2 + b^2 + \rho^2. \tag{12}$$

Furthermore, from (11) we can prove

$$\varepsilon^2 \geq \rho|\alpha(\theta)|. \tag{13}$$

Now, let us define $z := z_0 + \rho e^{i\theta}$. Then, from (11) we conclude that

$$z - \lambda_j = z_0 + \rho e^{i\theta} - \lambda_j = \frac{\lambda_k - \lambda_j}{2} + \rho e^{i\theta} = a + \rho\cos\theta + i\,(b + \rho\sin\theta)\,.$$

In a similar way, we have $z - \lambda_k = -a + \rho\cos\theta + i\,(-b + \rho\sin\theta)$. So, following from (11) and (12), we have

$$\begin{cases} |z - \lambda_j|^2 = a^2 + b^2 + \rho^2 + 2\rho\,(a\cos\theta + b\sin\theta)\,, \\ |z - \lambda_k|^2 = a^2 + b^2 + \rho^2 - 2\rho\,(a\cos\theta + b\sin\theta)\,, \end{cases}$$

that is,

$$\begin{cases} |z-\lambda_j|^2 = \varepsilon^2 + 2\alpha(\theta)\rho, \\ |z-\lambda_k|^2 = \varepsilon^2 - 2\alpha(\theta)\rho. \end{cases} \tag{14}$$

Now, let us suppose that $|z-\lambda_j| \leq |z-\lambda_k|$, by Proposition 14 we have $g^z_{jk} = g_{jk}(z, t^z_{jk})$, where t^z_{jk} is defined in (6). Now, by using (14), we have

$$t^z_{jk} = \frac{-2\alpha(\theta)\rho}{\varepsilon}. \tag{15}$$

On the other hand, following from expressions (8) and (9), as $g^z_{jk} = f^+_{z,j}(t^z_{jk})$ (Proposition 14), it is enough to see that $g^z_{jk} = f^+_{z,j}(t^z_{jk}) = \varepsilon$ for proving the lemma. Now, using the expression for $f^+_{z,j}(t^z_{jk})$ given in (5), we infer that

$$2(g^z_{jk})^2 = (t^z_{jk})^2 + 2|z-\lambda_j|^2 + t^z_{jk}\sqrt{(t^z_{jk})^2 + 4|z-\lambda_j|^2}.$$

Now, by using the expressions for $|z-\lambda_j|^2$ and t^z_{jk} derived from (14) and (15), respectively, we conclude that

$$\begin{aligned} (g^z_{jk})^2 &= \frac{2\alpha(\theta)^2\rho^2}{\varepsilon^2} + \varepsilon^2 + 2\alpha(\theta)\rho - \frac{2\alpha(\theta)\rho}{\varepsilon}\sqrt{\frac{\alpha(\theta)^2\rho^2}{\varepsilon^2} + \varepsilon^2 + 2\alpha(\theta)\rho} \\ &= \frac{2\alpha(\theta)^2\rho^2}{\varepsilon^2} + \varepsilon^2 + 2\alpha(\theta)\rho - \frac{2\alpha(\theta)\rho}{\varepsilon}\left|\frac{\alpha(\theta)\rho}{\varepsilon} + \varepsilon\right|. \end{aligned}$$

Now, in accordance with (13) it turns out that $\varepsilon + \alpha(\theta)\rho/\varepsilon \geq 0$, and we immediately have $(g^z_{jk})^2 = \varepsilon^2$.

The case $|z-\lambda_k| \leq |z-\lambda_j|$ can be reasoned in a similar way.

□

As a consequence of this theorem and Proposition 16, we infer the following result.

Corollary 18. *With the notations of Theorem* 17*,*
(a) *let z be such that $|z-z_0| = \rho_{jk}(\varepsilon)$ and*

$$\max\{|z-\lambda_j|, |z-\lambda_k|\} \leq |z-\lambda_i|, \quad i \neq j, k.$$

Then, $h_2(z) = \varepsilon$.
(b) *for every $z \in \mathcal{D}_{jk}(\varepsilon)$, where $\mathcal{D}_{jk}(\varepsilon)$ is defined in* (3)*, it is satisfied that $h_2(z) \leq \varepsilon$.*

We are in a position to prove Theorem 3.

Proof Theorem 3. First, let us supoose that $z \in \mathcal{D}_{jk}(\varepsilon)$. Then, by Corollary 18(b), it turns out that $h_2(z) \leq \varepsilon$. So, by Proposition 2 it is satisfied that $z \in \Lambda_{\varepsilon,2}(A)$.

Reciprocally, let $w \in \Lambda_{\varepsilon,2}(A)$, or, equivalently, $h_2(w) \leq \varepsilon$. Let us choose λ_j, λ_k such that

$$|w-\lambda_j| \leq |w-\lambda_k| \leq |w-\lambda_i|, \quad i \neq j, k.$$

Now, as in Theorem 17, we define $z_0 := \frac{\lambda_j+\lambda_k}{2}$, then we can write $w = z_0 + \rho_0 \mathrm{e}^{\mathrm{i}\theta}$, for certain $\rho_0 \geq 0$ and $\theta \in [0, 2\pi)$. Let us observe that by Corollary 18 (a), we have $h_2(w) = \varepsilon_0$, where

$$\varepsilon_0 := \sqrt{\rho_0^2 + \left(\frac{|\lambda_k - \lambda_j|}{2}\right)^2}.$$

So, as $h_2(w) \leq \varepsilon$, we conclude that

$$\sqrt{\rho_{jk}(\varepsilon)^2 + \left(\frac{|\lambda_k - \lambda_j|}{2}\right)^2} = \varepsilon \geq \varepsilon_0;$$

that is, $\rho_0 \leq \rho_{jk}(\varepsilon)$.

□

As a consequence of Theorem 3, we have the following result which describes the set $\Lambda_{\varepsilon,2}(A)$ when all eigenvalues of A are multiple.

Corollary 19. *Given a normal matrix $A \in \mathbb{C}^{n\times n}$ let us suppose that $\Lambda(A) = \Lambda_2^{(\mathrm{a})}(A) = \{\lambda_1, \ldots, \lambda_u\}$. Then, for every $\varepsilon \geq 0$ we have*

$$\Lambda_{\varepsilon,2}(A) = \bigcup_{i=1}^{u} \mathcal{D}(\lambda_i, \varepsilon).$$

Proof. Let us note that by Theorem 3, it is enough to prove that, for every pair of eigenvalues $\lambda_j \neq \lambda_k$ of A, it is satisfied

$$\mathcal{D}_{jk}(\varepsilon) \subset \mathcal{D}(\lambda_j, \varepsilon) \cup \mathcal{D}(\lambda_k, \varepsilon),$$

for each $\varepsilon \geq 0$. Let us observe that if $\varepsilon < \frac{|\lambda_j - \lambda_k|}{2}$, then $\mathcal{D}_{jk}(\varepsilon) = \emptyset$, that is to say, this case is trivial. So, let us suppose that $\varepsilon \geq \frac{|\lambda_j - \lambda_k|}{2}$. Now, if $z \in \mathcal{D}_{jk}(\varepsilon)$, following the notations of (11), from the proof of Theorem 17, we have

$$\begin{cases} |z - \lambda_j|^2 = \varepsilon^2 + 2\alpha(\theta)\rho, \\ |z - \lambda_k|^2 = \varepsilon^2 - 2\alpha(\theta)\rho. \end{cases}$$

So, if $|z - \lambda_j| \leq |z - \lambda_k|$, then $\alpha(\theta) \leq 0$ and $|z - \lambda_j| \leq \varepsilon$, therefore, $z \in \mathcal{D}(\lambda_j, \varepsilon)$. If $|z - \lambda_k| \leq |z - \lambda_j|$, it turns out that $\alpha(\theta) \geq 0$ and, therefore, $|z - \lambda_k| \leq \varepsilon$, so, $z \in \mathcal{D}(\lambda_k, \varepsilon)$.

□

We think it appropriate to make the following conjecture.

Conjecture 20. *If $A \in \mathbb{C}^{n\times n}$ is normal, then for $\varepsilon \geq 0$ and $k = 1, 2, \ldots, n$ the algebraic ε-pseudospectrum of order k of A is a union of closed disks, or the empty set.*

5 Geometric pseudospectra

Given a matrix $A \in \mathbb{C}^{n\times n}$, for $k = 1, 2, \ldots, n$ we write $\Lambda_k^{(\mathrm{g})}(A)$ for the set of eigenvalues of A of geometric multiplicity $\geq k$.

For each $\varepsilon \geq 0$, we define the *ε–pseudospectrum of geometric multiplicity $\geq k$* of A as the set

$$\Lambda_{\varepsilon,k}^{(\mathrm{g})}(A) := \bigcup_{\|X-A\|\leq\varepsilon} \Lambda_k^{(\mathrm{g})}(X).$$

Also, briefly we will call this pseudospectrum, the *geometric ε-pseudospectrum of order k* of A. Let us note that $\Lambda_{0,k}^{(\mathrm{g})}(A) = \Lambda_k^{(\mathrm{g})}(A)$. Let us also remark that for $\varepsilon \geq 0$,

$$\Lambda_{\varepsilon,2}^{(\mathrm{g})}(A) \subset \Lambda_{\varepsilon,2}^{(\mathrm{a})}(A).$$

Example 21. Let $A_0 \in \mathbb{C}^{4\times 4}$ the matrix

$$\mathrm{diag}(2, -1+\sqrt{3}\,\mathrm{i}, -1-\sqrt{3}\mathrm{i}, \frac{5}{2}+\frac{5}{2}\sqrt{3}\,\mathrm{i}).$$

In Figure 2 are showed the boundaries of $\Lambda_{\varepsilon,1}^{(\mathrm{g})}(A_0)$ in blue, $\Lambda_{\varepsilon,2}^{(\mathrm{g})}(A_0)$ in red, and $\Lambda_{\varepsilon,3}^{(\mathrm{g})}(A_0)$ in green, for $\varepsilon = 2.5$.

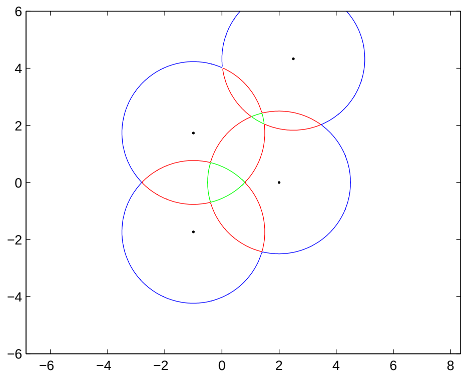

Figure 2: Boundaries of the geometric ε-pseudospectra.

We are going to describe the set $\Lambda_{\varepsilon,k}^{(\mathrm{g})}(A)$, when the matrix $A \in \mathbb{C}^{n\times n}$ is normal. Thus, denoting by

$$\mathcal{D}(\lambda, \delta) := \{z \in \mathbb{C} : |z - \lambda| \leq \delta\},$$

and by $\mathrm{Eig}(A)$ the n-tuple of unordered eigenvalues of A repeating the multiples, we arrive at the following theorem.

Theorem 22. *Let us assume that the matrix $A \in \mathbb{C}^{n\times n}$ is normal. Let us also suppose that* $\mathrm{Eig}(A) = (\lambda_1, \lambda_2, \ldots, \lambda_n)$*. Given $\varepsilon \geq 0$, for $k = 1, 2, \ldots, n$ we have*

$$\Lambda_{\varepsilon,k}^{(\mathrm{g})}(A) = \bigcup_{1\leq i_1<i_2<\cdots<i_k\leq n} \bigcap_{j=1}^{k} \mathcal{D}(\lambda_{i_j}, \varepsilon).$$

Consequently,

$$\Lambda^{(\mathrm{g})}_{\varepsilon,2}(A) = \bigcup_{1\le i<j\le n} \Big(\mathcal{D}(\lambda_i,\varepsilon)\cap\mathcal{D}(\lambda_j,\varepsilon)\Big).$$

Proof. As A is normal, we can assume without loss of generality that $A = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Hence, given $z \in \mathbb{C}$, the system of singular values of $zI_n - A$ are

$$|z-\lambda_1|, |z-\lambda_2|, \dots, |z-\lambda_n|,$$

not necessarily ordered. We are going to use the property

$$\Lambda^{(\mathrm{g})}_{\varepsilon,k}(A) = \{z \in \mathbb{C} : \sigma_{n-k+1}(zI_n - A) \le \varepsilon\}.$$

First, if $z \in \Lambda^{(\mathrm{g})}_{\varepsilon,k}(A)$, then $\sigma_{n-k+1}(zI_n - A) \le \varepsilon$, that is, there exist k terms of the unordered n-tuple $\operatorname{Eig}(A)$ that will be denoted by λ_{i_j}, $j = 1, 2, \dots, k$ such that

$$|z - \lambda_{i_j}| \le \varepsilon.$$

Therefore, $z \in \mathcal{D}(\lambda_{i_j}, \varepsilon)$ for $j = 1, 2, \dots, k$. The converse demonstration can be proved following the reverse reasoning.

□

6 Conclusions

If $A \in \mathbb{C}^{n\times n}$ is such that for every $\varepsilon > 0$ the ordinary ε-pseudospectrum of A is the union of the disks of radius ε centered at the eigenvalues of A, we have given a general proof that A must be a normal matrix.

The gist of the article is the following result. When A is a normal matrix, we have characterized its algebraic ε-pseudospectra of order 2 by means of a union of closed disks, ones centered at the multiple eigenvalues of A and radius ε, and the others centered at the midpoints of the line segments that join each pair λ_j, λ_k of different eigenvalues of A where at least one is simple, and radii $\rho_{jk}(\varepsilon) = \sqrt{\varepsilon^2 - \frac{|\lambda_j-\lambda_k|^2}{4}}$. Also, it is possible that for sufficiently small values of ε this pseudospectrum to be empty. A similar result for the geometric ε-pseudospectra of order k of a normal matrix A is given in the last section; this set is equal to the union of the $\binom{n}{k}$ intersections of k disks centered at the eigenvalues of A and radius ε. Besides, the set could be empty.

We are currently working on the Conjecture 20 for the $k = 3$ case, based on papers by Ikramov and Nazari, mainly [6].

References

[1] M. Aleksiejczyk, A. Smoktunowicz. On properties of quadratic matrices. *Mathematica Pannonica*, 11/2 (2000)239–248.

[2] G. Armentia, J.M. Gracia, F.E. Velasco. Derivatives of the diameter and the area of a connected component of the pseudospectra. Submitted to *Electronic J. Linear Algebra,* 2010.

[3] R. Bhatia. *Matrix analysis.* Springer, 1997.

[4] A. George, Kh. D. Ikramov. Unitary similarity of matrices with quadratic minimal polynomials. *Linear Algebra Appl.*, 349: (2002)11–16.

[5] A.S. Householder, J.A. Carpenter. The singular values of involutory and of idempotent matrices. *Numer. Math.* 5, (1963)234–237.

[6] Kh. D. Ikramov, A.M. Nazari. Normal matrices and an extension of Malyshev's formula. *Mathematical Notes* 75, Numbers 5-6, (2004) 608–616.

[7] M. Karow. *Geometry of spectral value sets.* Ph.D. Thesis. University of Bremen, 2003.

[8] A.N. Malyshev. A formula for the 2-norm distance from a matrix to the set of matrices with multiple eigenvalues. *Numer. Math.*, **83**, no. 3,(1999)443–454.

[9] L.N. Trefethen, M. Embree. *Spectra and pseudospectra. The behavior of nonnormal matrices and operators.* Princeton Univ. Press, 2005.

[10] F. Zhang. *Matrix theory. Basic results and techniques.* Springer, 1999.