Int. j. adv. multidisc. res. stud. 2022; 2(2):184-192

Received: 25-01-2022
Accepted: 05-03-2022

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

Bounds of Analytic Functions of Matrices in Banach Algebras

[1] **James Mbugua Ngigi**, [2] **Benard Okelo**

[1, 2] School of Biological, Physical, Mathematics and Actuarial Sciences, Jaramogi Oginga Odinga University of Science and Technology, Bondo-Kenya

Corresponding Author: **James Mbugua Ngigi**

Abstract

In this study we characterize bounds for analytic functions of matrices in Banach algebras. We consider both 2x2 and 3x3 matrices in Hilbert spaces and Banach algebras in general with Jordan canonical representation and upper triangular matrices.

Keywords: Analytic Functions, Banach Algebras, quantum physics

1. Introduction

An analytic function is a function that is locally given by a convergent power series [1]. There exist both real analytic functions and complex analytic functions [2]. These functions are infinitely differentiable. A complex analytic function is holomorphic i.e., it is complex differentiable. A matrix function is a function which maps a matrix to another matrix [3]. A matrix function can be lifted from a real function by power series, Jordan decomposition, Cauchy integral, matrix perturbations etc [4-8]. Bounds for analytic functions of matrices have attracted considerable interest over the years [9-11]. These bounds relate the size of the functions of matrices to the size of matrix polynomials over the numerical range [12-17]. The numerical range of a matrix is a convex subset of the complex plane, consisting of all Rayleigh quotients given by the following

$$W(A) = \left\{ \frac{x^* A x}{x^* x} \quad : x \neq 0, x \in \mathbb{C}^n \right\}$$

The numerical range has very nice properties including convexity and compactness [18]. Furthermore, the numerical range has applications in many areas including operator theory, dilation theory, C*-algebras and factorization of matrix polynomials (just to mention a few) and thus is a promising set to consider when seeking information about a matrix. The numerical range can also be generalized to higher rank numerical range, defined for $n \times n$ matrix A as

$$\Lambda_k(A) = \{\lambda \in \mathbb{C} : X^*AX = \lambda I_k \ for\ some\ n \times k\ matrix\ X\ with\ X^*X = I_k\},\ 1 \leq k \leq n.$$

This was first introduced by [19]. The higher rank numerical range has application in quantum error correction codes. The interest for the bounds of analytic function stems largely from its importance in a number of applications including numerical analysis, harmonic analysis, quantum chemistry, quantum information theory, random matrix theory, quantum physics, perturbation theory etc. In this work, we build on the work of [20] to investigate the bound for analytic function of 3×3 matrices with numerical range as ellipse [21]. We study the bound

$$\psi_\Omega(A) := sup\left\{\| f(A) \| : f \in H^\infty(\Omega), \| f \|_{l^\infty(\Omega)} \leq 1\right\}$$

And letting W(A) to be numerical range of A, we have

$$\psi(A) := sup\left\{\| P(A) \| : P\ polynomial, \| P \|_{l^\infty(W(A))} \leq 1\right\}$$

$= sup\{\psi_{\Omega}(A) : W(A) \subset \Omega\}$,

Where Ω is an open convex subset of the complex plane ($\emptyset \neq \Omega \neq \mathbb{C}$).

2. Preliminaries

In this section, we give some basic definitions, results and theorems used in this study.

Definition 1.2.1: A set $S \subset \mathbb{R}^n$ or $\mathbb{C}^n$ is said to be *convex* if the line segment connecting x and y is contained in S *i.e* $\forall x, y \in S \{ tx + (1-t)y : t \in [0,1] \} \subset S$.

Definition 1.2.2: Let $A \in \mathbb{C}^{n\times n}$. if $Ax = \lambda x$ for $\lambda \in \mathbb{C}, x \neq 0$ then λ is the *eigenvalue* of A and x is the *eigenvector* of A. The set of all eigenvalues of A is the *spectrum* of A denoted as $\sigma(A)$.

Definition 1.2.3: The *convex hull* of a set S, denoted $co(S)$, is the minimal convex set containing S.

Definition 1.2.4: For any analytic function f defined on a set $D \subset \mathbb{C}$, it holds

$$f(x) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z-x} dz$$

where C is a closed curve inside the domain D enclosing x.This is called the *Cauchy integral formula.*

Definition 1.2.5: Let $T : H_1 \to H_2$ be a bounded linear operator, where H_1 and H_2 are Hilbert spaces.Then the Hilbert *adjoint operator* T^* of T is the operator $T^* : H_2 \to H_1$ such that for all $x \in H_1$ and $y \in H_2$, $\langle Tx, y\rangle = \langle x, T^*y\rangle$.

Definition 1.2.6: Two vectors x and y in an inner product space E are said to be orthogonal if and only if $\langle x, y\rangle = 0, x \neq y$, denoted by $x^{\perp}y$.

Definition 1.2.7: Let S be a subset of a Hilbert space H. The set of all vectors orthogonal to S is called the *orthogonal complement* of S denoted by $S^{\perp}$.i.e., $S^{\perp} = \{x \in H : x^{\perp}s, \forall s \in S\}$.

Definition 1.2.8: Let $A \in M_n$ .Write $A = H + iK$ with H, K Hermitiian, and let $L_A(u, v, w) = det(uH + vK + wI)$. The equation $L_A(u, v, w) = 0$, with u, v, w viewed as homogeneous line coordinates defines an *algebraic curve* of class n called *Kippenhahn polynomial*.

Definition 1.2.9: The real part of the algebraic curve $L_A(u, v, w) = 0$ is called the *associated curve* denoted by $C(A)$.

Definition 1.2.10: We say that a *matrix* A *is reducible* if there exist a unitary matrix U such that

$$U^*AU = diag[A_1, A_2] \quad ,$$

where both diagonal blocks are of non-zero size.

Definition 1.2.11: Let H be a Hilbert space. An operator $T : H \to H$ is called *Hermitian or self-adjoint* if $T = T^*$.

Definition 1.2.12: Let H be a Hilbert space. An operator $T : H \to H$ is called *unitary* if $TT^* = T^*T = I$.

Definition 1.2.13: A complex valued function h of a complex variable λ is said to be *holomorphic* (or analytic) on a domain G of the complex $\lambda -$plane if h is defined and differentiable on G . That is, the derivative h' of h defined by

$$h'(\lambda) = \lim_{\Delta\lambda} \frac{h(\lambda+\Delta\lambda)-h(\lambda)}{\Delta\lambda}$$ exists for every $\lambda \in G$

Definition 1.2.14: Let $a \in \bar{D}(0,1)$, the unit disk centred at origin. Then the *Blaschke factor* is defined by

$$B_a(z) = \frac{z-a}{1-\bar{a}z}$$

Definition 1.2.15. A *Blaschke product* is an expression of the form

$$B(z) = z^m \prod_{j=1}^{\infty} \frac{-\overline{a_j}}{\|a_j\|} Ba_j(z)$$

, where m is non-negative integer.

Definition 1.2.16: (Young's inequality) If a and b are non-negative real numbers and p and q are positive real numbers such

that $\frac{1}{p} + \frac{1}{q} = 1$,

then $ab \le \frac{a^p}{p} + \frac{b^q}{q}$. When $p = q = 2$, we have $ab \le \frac{a^2}{2} + \frac{b^2}{2}$. Equality occurs if and only if $b = a^{p-1}$.

Definition 1.2.17: (Nevanlinna-Pick interpolation problem) In complex analysis, given initial data consisting of n points $\lambda_1, \lambda_2, \ldots, \lambda_n$ in the complex unit disc D and target data consisting of n points $z_1, z_2, \ldots, z_n$ in D , the *Nevanlinna –Pick interpolation problem* seeks to find a holomorphic function φ that interpolates the data ,that is for all i,

$$\varphi(i) = z_i$$

subject to the constrain

$$|\varphi(\lambda)| \le 1$$

For all $\lambda \in D$.

Definition 2.0.1: The numerical range of $A \in M_n$, is the subset $W(A) \subset \mathbb{C}$, given by

$$W(A) = \{\langle Ax, x\rangle : x \in \mathbb{C}^n; \|x\| = 1\},$$

where $\|.\|$ denotes the 2-norm. Note that $W(A)$ is the continuous image of a compact set, and is thus itself a compact set in $\mathbb{C}$. As we will show, the numerical range of a linear operator is a convex set. This is a consequence of the *Toeplitz-Hausdorff Theorem*. We first review some basic properties of the numerical range.

3. Main results

In this section, we give the results of our study. We begin with the following proposition. Let Ω be a convex subset of the complex plane; we assume $\Omega \neq \emptyset, \Omega \neq \mathbb{C}$.Since Ω is simply connected there exists a holomorphic bijection ϕ from Ω onto the open unit disk

$$D = \{z \in \mathbb{C} : |z| < 1\}.$$

From the Osgood-Caratheodory Theorem ϕ admits an extension which is a homeomorphism from $\overline{\Omega}$ onto $\overline{D}$. It is convenient to introduce the Blaschke functions

$$b_\xi = \frac{\phi(\xi) - \phi(z)}{1 - \overline{\phi(\xi)}\phi(z)}, \quad \xi \in \Omega \tag{3.1.1}$$

It is clear $b_\xi \in H^\infty(\Omega) \cap C^0(\Omega), \quad \|b_\xi\|_{L^\infty}(C) = 1,$ and if $z \in \partial\Omega$ then $|b_\xi(z)| = 1.$

We introduce also the set of finite Blaschke products

$$\beta_k(\Omega) = \{f : f(z) = e^{i\varphi} \textstyle\prod_{j=1}^{r} b_{\xi_j}(z), \quad \varphi \in \boldsymbol{R}, \quad \xi_j \in \Omega, \ j = 1, \ldots, r \ , \ 0 \le r \le k\})$$

We use the convention $f(z) = e^{i\varphi}$ if $r = 0$ so $\beta_{(0)}(\Omega)$ corresponds to the constant function of modulus 1. It can be seen that this space is independent of the choice of the isomorphism ϕ between Ω and D. The following theorem is a consequence of the Nevanlinna-Pick theory.

Theorem 3.1.0: Let $A \in \mathbb{C}^{d,d}$ be a square matrix with $\sigma(A) \subset \Omega$. Then there exist a function $f \in H^\infty(\Omega)$ such that

$\|f(A)\| = \psi_{(\Omega)}(A)$ and $\|f\|_{L^\infty(\Omega)} = 1$. Any such a function belongs to $\beta_{d-1}(\Omega)$.

Proof: Let $\{\lambda_j\}$, $j = 1, \dots, k$ be the set of the distinct eigenvalues of A and $\{r^i\}$ the corresponding multiplicities. Writing A in Jordan form we can see that $f(A)$ only depends on A and the values $\{f^{(l)}(\lambda_j), 0 \le l < r_j, 1 \le j \le k\}$.
From the compactness property of $H^\infty(\Omega)$ we deduce that there exists $f \in H^\infty(\Omega)$ such that $\|f\|_{l^\infty(\Omega)} = 1$ and $\|f(A)\| = \psi_{(\Omega)}(A)$.
Then let g be the Nevanlinna–Pick interpolant of f, i.e g is the function which minimizes $\|g\|_{l^\infty(\Omega)}$ among the solutions of

$$g \in H^\infty(\Omega), g^{(l)}(\lambda_j) = f^{(l)}(\lambda_j), 0 \le l < r_j, 1 \le j \le k. \quad (3.2.2)$$

It is known that there exists $c \in \boldsymbol{R}^*$ such that $g/c \in \beta_{d-1}(\Omega)$: We have

$$c = \|g\|_{l^\infty(\Omega)} \le \|f\|_{l^\infty(\Omega)} = 1$$

and g is unique. But 3. 2.2 implies $(A) = f(A)$, therefore

$$\psi_{(\Omega)}(A) = \|g(A)\| = \psi_{(\Omega)}(A)/\|g\|_{l^\infty(\Omega)}.$$

We deduce $\|g\|_{l^\infty(\Omega)}$ =1 and $g = f$.

Remark 3.1.1: We remark on some properties of the function $\psi_{(\Omega)}(A)$.
a) If ϕ denotes a holomorphic isomorphism from Ω to the unit disc D we have $\psi_{(\Omega)}(A) = \psi_{(D)}(\phi(A))$. Indeed $f(A) = g(\phi(A))$ with $g := f o \phi^{-1}$, and clearly f and g have the same maximum norm.
b) If $A = H^{-1}TH$ we have $f(A) = H^{-1}f(T)H$, thus $\psi_{(\Omega)}(A) = \psi_{(\Omega)}(T)$ as soon as the matrix is unitary. That allows us to restrict our study to the case of upper triangular matrices without loss of generality.

Lemma 3.1.1: The bound $\psi_{(\Omega)}(A)$ depends continuously on A and is decreasing with respect to Ω (strictly decreasing if $\psi_{(\Omega)}(A) \ne 1$). Furthermore

$$\psi(A) = \sup\{\psi_{(\Omega)}(A): W(A) \subset \Omega\}. \quad (3.3.3)$$

Proof: (a) The continuity with respect to A follows from the Cauchy integral representation

$$f(A) - f(A') = \frac{1}{2\pi i}\oint_c f(z)((z-A)^{-1} - (z-A')^{-1})\, dz,$$

where c is an oriented curve surrounding the spectrum of A (and therefore of A' for A' close enough to A), we deduce that $f(A')$ tends to $f(A)$ as $A' \to A$ uniformly with respect to the function f bounded by 1.

b) If we have $\sigma(A) \subset \Omega \subseteq \Omega'$, then we have

$$\psi_{(\Omega)}(A) > \psi_{(\Omega')}(A).$$

Indeed let $f \in \beta_{d-1}(\Omega')$ such that $\|f(A)\| = \psi_{(\Omega')}(A)$ and $\|f\|_{l^\infty(\Omega')} = 1$,we have clearly $\|f(A)\| \le \psi_{(\Omega)}(A)$ since $\|f\|_{l^\infty(\Omega')} = 1$, and we cannot have $\|f(A)\| = \psi_{(\Omega)}(A)$ except if

$$f \in \beta_{d-1}(\Omega) \cap \beta_{d-1}(\Omega'), \text{ i.e., if } f \text{ is constant.}$$

Note that this induces a continuity with respect to Ω. Indeed, we can assume that $0 \in \Omega$ without loss of generality. Then we set, for $x > 0$, $\Omega_x = x\Omega$, and we clearly have $\psi_{\Omega_x}(A) = \psi_\Omega(x^{-1}A)$, from part (a) $\psi_{\Omega_x}(A)$ continuously depend on x. If Ω' is another convex set close to Ω we have

$$\left| \psi_{(\Omega)}(A) - \psi_{(\Omega')}(A) \right| \le \psi_{\Omega_x}(A) - \psi_{\Omega_y}(A), \text{ if } \Omega_x \subset \Omega' \subset \Omega_y \text{ and } x < 1 < y.$$

That implies continuity with respect to Ω for the Hausdorff distance.
c) Turning now to the proof of (3.3.3), we first look at the case when the matrix A is normal. Then we have $\psi(A) = \psi_{(\Omega)}(A) = 1$, $\forall\, \Omega \supset \sigma(A)$ and the result is straight forward.

In the other cases the interior of $W(A)$ is not empty and we set $\Omega = int\, W(A)$. If $\sigma(A) \subset \Omega$ then we have clearly $\psi(A) \leq \psi_{(\Omega)}(A)$ and from the previous theorem there exist a function

$f \in \beta_{d-1}(\Omega)$ such that $\|f(A)\| = \psi_{(\Omega)}(A)$. But $\beta_{d-1}(\Omega) \subset C^0\overline{(\Omega)}$, thus we can find a sequence of polynomials p_n which uniformly tends to f in Ω which proves that $\psi_{(\Omega)}(A) \leq \psi(A)$ .When

$\sigma(A) \cap \partial\Omega \neq 0,$ we can find a unitary matrix U such that

$$A = U^* \begin{pmatrix} A_2 & 0 \\ 0 & A_3 \end{pmatrix} U$$

with A_2 diagonal and $\sigma(A_1) \subset \Omega$ .Then it is clear that $\psi(A) = \psi(A_1)$ and for all $\Omega' \supset \sigma(A_1), \psi_{(\Omega')}(A) = \psi_{\Omega'}(A_1)$. Therefore, the result follows from the previous case.

Now we give an explicit formula in the case of 2×2 matrix. In order to express it we introduce the following function.

$$\theta(x,y) := \frac{x + \sqrt{1 + x^2}}{y + \sqrt{1 + y^2}}$$

Proposition 3.1.3: If

$$T = \begin{pmatrix} \lambda_1 & \gamma \\ 0 & \lambda_2 \end{pmatrix},$$

then we have

$$\psi_D(T) = max\left(1, \theta\left(\frac{|\gamma|}{|\lambda_1-\lambda_2|}, \frac{\sqrt{(1-|\lambda_1|^2)(1-|\lambda_2|^2}}{|\lambda_1-\lambda_2|} \right)\right), \quad \text{if } \lambda_1 \neq \lambda_2$$

$$\psi_D(T) = \max\left(1, \frac{|\gamma|}{1-|\lambda_2|}\right), \quad \text{if } \lambda_1 = \lambda_2$$

Proof: a) By continuity it is sufficient to consider the case $\lambda_1 \neq \lambda_2$. We define

$$\Psi_1(T) = \frac{|\gamma|}{|\lambda_1-\lambda_2|},$$

$$\Psi_2(T) = \frac{\sqrt{(1-|\lambda_1|^2)(1-|\lambda_2|^2}}{|\lambda_1-\lambda_2|},$$

And we recall that

$$a(T) = \begin{pmatrix} \phi(\lambda_1) & \gamma\phi(\lambda_1,\lambda_2) \\ 0 & \phi(\lambda_1) \end{pmatrix}$$

Where,

$$\phi(\lambda_1, \lambda_2) = \frac{\phi(\lambda_1)-\phi(\lambda_2)}{\lambda_1-\lambda_2} \tag{3.3.4}$$

It is easily verified that if ϕ is an automorphism of D the three quantities $\psi_D(T), \Psi_1(T), \Psi_2(T)$ remain invariant if we replace the matrix T by $\phi(T)$.Indeed we have just to verify this for automorphisms of the form $\phi(z) = e^{i\varphi}z, \varphi \in \boldsymbol{R}$ and $\phi(z) = \frac{(x-z)}{(1+xz)}$, $x \in [0,1]$. Since such mappings ϕ generate the automorphism group of D. Note also that $\psi_D(T)$ does not change if we replace γ by $|\gamma|$ since ψ_D is invariant by a unitary similarity. For all λ_1 and $\lambda_2 \in D$ it is possible to find an automorphism ϕ such that $\phi(\lambda_1) + \phi(\lambda_1) = 0$ and $\phi(\lambda_1) \in (0,1)$.Therefore it is sufficient to prove the proposition in the case where the matrix T is of the form.

$$T = \begin{pmatrix} \lambda & 2\delta \\ 0 & -\lambda \end{pmatrix}, \lambda \in (0,1), \delta \geq 0 \qquad (3.3.5)$$

b) We now consider this case 3.3.5. A simple computation shows that $\|T\| = \delta + \sqrt{\lambda^2 + \delta^2}$ and in this situation the statement of the theorem reads.

$$\psi_D(T) = \max(1, \|T\|).$$

If $\|T\| \leq 1$, a well- known von Neumann inequality asserts that $\psi_D(T) = 1$, thus we only have to consider the case $\|T\| > 1$. It is clear that $\psi_D(T) \geq \|T\|$ (take $f(z) = z$ in the definition of $\psi_D(T)$). For the converse inequality we set

$$\beta = \frac{1+\lambda^2}{2\sqrt{(\lambda^2+\delta^2)}}, \mu = \frac{1-\lambda^2-2\beta\delta}{2\lambda}, H = \begin{pmatrix} 1 & \mu \\ 0 & \beta \end{pmatrix};$$

Then we have

$$B = H^{-1}TH = \begin{pmatrix} \lambda & 2(\lambda\mu + \beta\delta) \\ 0 & -\lambda \end{pmatrix} = \begin{pmatrix} \lambda & 1-\lambda^2 \\ 0 & -\lambda \end{pmatrix}.$$

This matrix satisfies $\|B\| = 1$, thus $\psi_D(B) = 1$, and consequently

$$\psi_D(T) \leq \|H\| \, \psi_D(B) \|H^{-1}\|.$$

But this quantity is the largest root of the equation

$$x^2 - \frac{1+\mu^2+\beta^2}{\beta}x + 1 = 0 \qquad (3.3.6)$$

If we define $x_1 = \sqrt{\delta^2 + \lambda^2 + \delta}$, then we have $x_2 = 1/x_1$ and by a simple computation

$$\beta(x_1 + x_2) - \mu^2 - \beta^2 = 1$$

We deduce that x_1 and x_2 are the two roots of 3.6, thus $\|H\| \ \|H^{-1}\| = 1 = \|T\|$, which implies the proposition.

Corollary 3.1.4: If $T = \begin{pmatrix} \lambda_1 & \gamma \\ 0 & \lambda_1 \end{pmatrix}$, and if Φ denotes a holomorphic bijection from Ω onto the unit disk D, then

$$\psi_\Omega(T) = max\left(1, \theta\left(\frac{|\gamma|}{|\lambda_1-\lambda_2|}, \frac{\sqrt{(1-|\lambda_1|^2)(1-|\lambda_2|^2}}{|\lambda_1-\lambda_2|} \right)\right), \qquad \text{if } \lambda_1 \neq \lambda_2$$

and

$$\psi_\Omega(T) = \max\left(1, \frac{|\gamma|}{1-|\lambda_2|}\right), \qquad \text{if } \lambda_1 = \lambda_2$$

Proof: From remark 3.1.1(a) we have $\psi_{(\Omega)}(T) = \psi_{(D)}(a(T))$. We apply the previous proposition and use formula 3.3.4.

Now we consider **Bound** for 3×3 matrices. We begin with the following classification concerning the numerical range of a 3×3 matrix A based on factorability of L_A, given by Kippenhahn.

Case 1: L_A factors into three linear factors. Then $C(A)$ consists of three (not necessarily distinct) points. A is normal (and therefore reducible), and $W(A)$ is the convex hull of its eigenvalues.

Case 2: L_A factors into a linear factor and a quadratic factor. Then $C(A)$ consists of a point λ_0 (the eigenvalue of A corresponding to the linear factor) and an ellipse E.The numerical range is either an ellipse (if λ_0 lies inside E) or a "cone like" figure otherwise; in the latter case A is reducible (but not normal).

Case 3: L_A is irreducible and the degree of $C(A)$ equals 4. Then $C(A)$ has a double tangent and the boundary of $W(A)$ contain one flat portion but no angular points.
Case 4: L_A is irreducible and the degree of $C(A)$ equals 6. Then $C(A)$ consists of two parts one inside another; an outer part (and therefore $W(A)$) has an ovular shape.
We consider case 2 of the Kippenhahn classification in which $W(A)$ is an ellipse.The following results are known.

Theorem 3.2.1: Let A be an $n \times n$ matrix with eigen values and suppose that its associated curve $C(A)$ consists of k ellipses, with minor axes of lengths $s_1, s_1, s_2, \dots, s_k$,and $n-2k$ points. Then

$$\sum_{i=0}^{k} s_i^2 = tr(A^*A) - \sum_{i=0}^{k} |\lambda_i|^2 \tag{3.2.1}$$

For $n = 3$ conditions of theorem 3.2.1 are satisfied for $C(A)$ being an ellipse and a point and in this case equation 3.2.1 takes the form

$$s = (tr(A^*A) - |\lambda_I|^2 - |\lambda_2|^2 - |\lambda_3|^2)^{1/2} \tag{3.2.2}$$

Proof: Relabel the eigenvalues of A in such a way that $\lambda_{2i-1}, \lambda_2$ become the foci of the i^{th} ellipse $(i = 1, \dots, k)$ and the remaining points of $C(A)$. Along with A, consider the matrix

$$B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \oplus \begin{pmatrix} \lambda_3 & s_1 \\ 0 & \lambda_4 \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} \lambda_{2k-1} & s_2 \\ 0 & \lambda_{2k} \end{pmatrix} \oplus diag[\lambda_{2k+1}, \dots, \lambda_n]$$

Since $C(A) = C(B)$, the polynomials L_A and L_B have to be the same. Compute now the coefficients of w^{n-2} of these polynomials. When doing that, due to unitary invariance of L_A, we may without loss of generality suppose that $A = [a_{ij}]_{i,j=1}^{n}$ is in upper –triangular form. The coefficient of w^{n-2} in L_A equals the sum of all 2×2 principal minors of $uH + vK$, that is,

$$\sum_{1 \le i < j \le n} \left[(u\Re a_{ii} + v\Im a_{ii})(u\Re a_{jj} + v\Im a_{jj}) - \frac{1}{4}(u^2+v^2)|a_{ij}|^2 \right]$$

$$= \sum_{1 \le i < j \le n} \left[(u\Re\lambda_i + v\Im\lambda_i)(u\Re\lambda_j + v\Im\lambda_j) - \frac{1}{4}(u^2+v^2)|a_{ij}|^2 \right]$$

Appling this formula to B (which already is in upper triangular form) we obtain

$$\sum_{1 \le i < j \le n} \left[(u\Re\lambda_i + v\Im\lambda_i)(u\Re\lambda_j + v\Im\lambda_j) - \frac{1}{4}(u^2+v^2)\sum_{i=1}^{n}|s_i|^2 \right]$$

Since L_A=L_B, it follows from here that

$$\sum_{i=1}^{n}|s_i|^2 = \sum_{1 \le i < j \le n}|a_{ij}|^2 = \sum_{i,j=1}|a_{ij}|^2 - \sum_{i=1}^{n}|a_{ii}|^2 = tr(A^*A) - \sum_{i=1}^{n}|\lambda_i|^2.$$

Note that in setting of Theorem 3.2.1 all the respective coefficients of L_A and L_B are equal. In particular, equating the coefficients of u^n, v^n yields

$$\det H = \prod_{i=1}^{k}\left(\Re\lambda_{2i-1}\Re\lambda_{2i} - \tfrac{1}{4}s_i^2\right)\prod_{i=2k+1}^{n}(\Re\lambda_i), \tag{3.2.3}$$

$$\det K = \prod_{i=1}^{k}\left(\Im\lambda_{2i-1}\Im\lambda_{2i} - \tfrac{1}{4}s_i^2\right)\prod_{i=2k+1}^{n}(\Im\lambda_i).$$

If $n = 3$ and A is in upper triangular of the form

$$A = \begin{pmatrix} a & x & y \\ 0 & b & z \\ 0 & 0 & c \end{pmatrix} \tag{3.2.4}$$

Condition 3.2.3 can be rewritten as $|x|^2\Re c + |y|^2\Re b + |z|^2\Re a - \Re(x\bar{y}z) = s^2\Re\lambda_3$,

$|x|^2\Im c + |y|^2\Im b + |z|^2\Im a - \Im(x\bar{y}z) = s^2\Im\lambda_3$, or simply

$$|x|^2 c + |y|^2 b + |z|^2 a - (x\bar{y}z) = s^2\lambda_3. \quad (3.2.5)$$

Due to equation 3.2.2,

$$s = \sqrt{|x|^2 + |y|^2 + |z|^2} \quad (3.2.6)$$

Hence the conditions 3.2.5, 3.2.6 are necessary for matrices 3.2.4 and

$$B = \begin{pmatrix} \lambda_1 & s & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

to have the same associated curves. Therefore, the following criterion holds.

Theorem 3.2.2: Let A be in upper triangular form 3.2.4. Then its associated curve $C(A)$ consists of an ellipse (possibly degenerating to a disk) and a point iff

1. $d = |x|^2 + |y|^2 + |z|^2 > 0$ and
2. The number $\lambda = (|x|^2 c + |y|^2 b + |z|^2 a - (x\bar{y}z))/d$ coincides with at least one of the eigenvalues a, b, c.

If these conditions are satisfied, then $C(A)$ is the union of λ with the ellipse having its foci at two other eigenvalues of A and minor axis of length $s = \sqrt{d}$.

Theorem 3.2.3: Let A be a 3×3 matrix with the eigenvalues λ_j, $j = 1,2,3$. Then $W(A)$ is an ellipse iff conditions 1,2 of theorem 3.2.2 hold and in addition,

3. $(|\lambda_1 - \lambda_3| + |\lambda_2 - \lambda_3|)^2 - |\lambda_1 - \lambda_2|^2 \leq d$, where the eigenvalue coinciding with λ is labeled λ_3 .

Proof: Conditions 1, 2 are equivalent to $C(A)$ being a union of the ellipse (with foci at λ_1, λ_2 and minor axis of length) and the point λ_3 .Condition 3 means that lies inside. According to Kippenhahn's classification, this is the only case when $W(A)$ is an ellipse.

Proposition 3.2.4:

Let $A = \begin{pmatrix} a & x & y \\ 0 & b & z \\ 0 & 0 & c \end{pmatrix}$ satisfying conditions of theorems 3.2.2 and 3.2.3. then we have

$$\psi_D(A) = max\left(1, \theta\left(\frac{\sqrt{d}}{|\lambda_1 - \lambda_2|}, \frac{\sqrt{(1-|\lambda_1|^2)(1-|\lambda_2|^2}}{|\lambda_1 - \lambda_2|} \right)\right) \qquad \text{if } \lambda_1 \neq \lambda_2$$

$$= max\left(1, \theta\left(\frac{\sqrt{|x|^2+|y|^2+|z|^2}}{|\lambda_1 - \lambda_2|}, \frac{\sqrt{(1-|\lambda_1|^2)(1-|\lambda_2|^2)}}{|\lambda_1 - \lambda_2|} \right)\right) \qquad \text{if } \lambda_1 \neq \lambda_2$$

And

$$\psi_D(A) = \max\left(1, \frac{\sqrt{d}}{1-|\lambda_2|}\right), \qquad \text{if } \lambda_1 = \lambda_2$$

$$= \max\left(1, \frac{\sqrt{|x|^2+|y|^2+|z|^2}}{1-|\lambda_2|}\right) \qquad \text{if } \lambda_1 = \lambda_2$$

4. Conclusion

In this paper, we have obtained an explicit formula for the bound $\psi_D(A)$ where A is a 3×3 matrix with elliptical numerical range. This has been done by reducing it to the case of 2×2 matrices whose numerical range is elliptical. These results can be used to obtain bounds for matrices with Jordan canonical representation and upper triangular matrices. This applies strictly to functions of 3×3 matrices which include a wide variety of functions arising in Mathematical Physics, numerical analysis, network science etc.

5. References

1. Bonsall FF, Duncan J. Numerical Ranges of Operator on Normed Spaces and of Elements of Normed Algebras, Cambridge university press, 1971. ISBN 978-0-521-07988-4
2. Bernard D, Francis D. Generalization of von Neumann's spectral sets and integral representation of operators. Bull. Soc. Math. France. 1999; 127(1):25-41.
3. Badea C, Crouzeix M, Delyon B. Convex Domains and k-spectral Sets. Math. Z. 2006; 252(2):345-365.
4. Cowen C, Harel E. An effective algorithm for computing the numerical range, August 1995.
5. Keeler DS, Rodman L, Spitkovsky IM. The numerical range of 3×3 matrices, Linear algebra and its applications. 1997; 252:115-139.
6. Hwa LG, Wu PY. Higher Rank numerical Ranges and Kippenhahn Polynomials, linear algebra and its applications. 2013; 438:3054-3061.
7. Shapiro J. Notes on numerical range, 2003. http:/www.math.msu.edu/~shapiro/pubvit/downloads.numrangenotes/numrange-notes-pdf
8. Akaashueek M, Ball JA, Gohberg I. Nevanlinna-Pick interpolation for time-varying input-output maps. The continuous time case, sc, Birkhausser, 1992.
9. Gustafson K, Duggirala K, Rao M. Numerical range, Universitext. Springer-Verlag. New York, 1997.
10. Luo SM. Crouzeix conjecture and the GMRES algorithm, All Thesis and Dissertations, 2011, article-2819.
11. Crouzeix M. Bounds for Analytic Functions of Matrices. Integral equations operator theory. 2004; 48(4):461-477.
12. Choi MD, Kribs DW, Zyczkowsi K. Higher Rank Numerical Ranges and Compression Problems, Linear algebra and its applications. 2006; 418:828-839.
13. Miguel M, Javier M, Rafael P. On the intrinsic and the spatial numerical range, Mathematical analysis and applications. 2006; 318:175-189.
14. Crouzeix M. Numerical Range and Functional Calculus in Hilbert Space. Journal of Functional Analysis. 2007; 244(2):668-690.
15. Fieldler M. Geometry of the Numerical Range of Matrices, Linear algebra and its applications. 1981; 37:81-96.
16. Crouzeix M, Delyon B. Some estimates for analytic functions of strip or sectorial operators. Arch. Math. (Basel). 2003; 81(5):559-566.
17. Halmos PR. A Hilbert space problem book, D.Van Nostrand Co. Inc. Princeton, N.J-Toronto, Ont.-London, 1967.
18. Panayiotis JP, Tsatsomeros MJ. Numerical range :(in) a matrix nutshell, May 2002.
19. Ando T. Structure of operators with numerical radius one, Acta Sci.Math. (Szeged). 1973; 34:11-15.
20. Tosio K. Some mapping theorems for the numerical range. Proc. Japan Acad. 1965; 41:652-655.
21. Young WH. On classes of summable functions and their Fourier series. Proceedings of the royal society. 1912; 87(594):225-229.