BOUNDS FOR ANALYTIC FUNCTIONS OF MATRICES

Michel CROUZEIX

If f is an analytic function bounded on a convex domain of the complex plane and A a square matrix whose spectrum is included in this domain, the function $f(A)$ is well defined. In this paper we study bounds for $\|f(A)\|$ uniform with respect to the functions f bounded by 1, and uniform with respect to the matrices A whose the numerical ranges are included in the domain. We show that these bounds are attained and give explicit formulae in some 2-dimensional cases.

1 Introduction

In all this paper we are working in an Hilbertian context. We use the notations (for generic elements $x,\, y \in \mathbb{C}^d$, square matrix $A \in \mathbb{C}^{d,d}$) :

$\langle x, y\rangle = y^* x = \sum_{j=1}^d x_j\,\overline{y_j}$ is the usual inner product in $\mathbb{C}^d$,

$\|x\| = \langle x, x\rangle^{1/2}$ is the Euclidian norm in $\mathbb{C}^d$,

$\|A\| = \max\{\|Ax\|\,;\, x \in \mathbb{C}^d,\ \|x\| \leq 1\}$ is the operator norm.

Let Ω be an open (non void) convex subset of the complex plane and A be a square matrix with spectrum $Sp(A) \subset \Omega$. We first consider the bound

$$\psi_\Omega(A) := \sup\{\|f(A)\|\,;\, f \in \mathcal{H}^\infty(\Omega),\ \|f\|_{L^\infty(\Omega)} \leq 1\}, \tag{1}$$

where $\mathcal{H}^\infty(\Omega)$ denotes the Hardy space

$$\mathcal{H}^\infty(\Omega) := \{f\,;\, f \text{ holomorphic and bounded in } \Omega\}.$$

Let $W(A)$ be the numerical range of A

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

we also consider the bound

$$\psi(A) := \sup\{\|p(A)\|\,;\ p \text{ polynomial},\ \|p\|_{L^\infty(W(A))} \leq 1\}. \tag{2}$$

These first two bounds are related to a family of analytic functions but to only one matrix. Afterwards we consider a family of matrices

$$\mathcal{M}_d(\Omega) = \{A \in \mathbb{C}^{d,d}\,;\, W(A) \subset \overline{\Omega} \text{ and } Sp(A) \subset \Omega\},$$

and the bound

$$\begin{aligned} C_\Omega(d) &:= \sup\{\|f(A)\|\,;A\in\mathcal{M}_d(\Omega),\ f\in\mathcal{H}^\infty(\Omega),\ \|f\|_{L^\infty(\Omega)}\le 1\},\\ &= \sup\{\psi_\Omega(A)\,;A\in\mathcal{M}_d(\Omega)\}. \end{aligned} \tag{3}$$

It is clear that $C_\Omega(d)$ only depends on d and on the shape of Ω (more precisely $C_\Omega(d) = C_{s(\Omega)}(d)$ if s is a similarity : $s(z) = a + b\,z$, $a, b \in \mathbb{C}$, $b \neq 0$, or an anti-similarity : $s(z) = a + b\,\bar{z}$). A theorem due to J. von Neumann [11] asserts that $C_\Omega(d) = 1$ when Ω is a half-plane. In the reference [5] it is proved that $C_\Omega(d)$ is uniformly bounded in d when Ω is convex and bounded, $C_\Omega(d) \le 3$ if Ω is a disk (see also [3]), and in [4] it is proved that $C_\Omega(d) \le 2 + 2/\sqrt{3}$ if Ω is a strip or a convex sector of the complex plane.

The organization of the paper is the following. In Section 2 we prove that the bound $\psi_\Omega(A)$ is attained by a function which is a Blaschke product, we give some properties of this bound and an explicit formula in the case of 2×2 matrices. In this last case the numerical range is an ellipse and we express $\psi(A)$ as a function of the ellipticity. In particular we show that $\psi(A) \le 2$ which implies the following result

THEOREM 1. *Let A be a 2×2 matrix with entries in $\mathbb{C}$. Then for all polynomials p we have*

$$\|p(A)\| \le 2 \max_{z\in W(A)} |p(z)|.$$

In Section 3 we show that the bound $C_\Omega(d)$ is attained by a (upper triangular) matrix $A \in \mathcal{M}_d(\Omega)$ and a Blaschke product f in the three following cases : Ω is bounded and its boundary is analytic, Ω is a convex sector or Ω is a strip. We prove that the strip case constant is the limit of the sector constant as the aperture tends to 0 and include the result that $C_\Omega(d) = 2$ if Ω is a disk (this result has been pointed to me by Catalin Badea [2], answering to a question posed in a previous version of this paper). Section 4 is devoted to the 2×2 case. Since $\psi(A) \le 2$ we have $C_\Omega(2) \le 2$, we show that $C_\Omega(2) = 2$ implies that Ω is a disk. Finally we give (quasi)-explicit formulae for $C_\Omega(2)$ when Ω is a sector or a strip of the complex plane.

CONJECTURE. I have not been able to answer to the following questions :

- is $C_\Omega(d)$ depending on d (for $d \ge 2$) ?
- in the two dimensional case we have proved that the disk is corresponding to the worst constant, does this result remain true in any dimension ?

I conjecture a positive answer to the last question. That would imply that Theorem 1 would generalize to any dimension, which will be a very nice result. I end this section by a last minor question :

- in the case "Ω is an ellipse of excentricity ε", is the constant $C_\Omega(2)$ equal to the right handside of equation (9) ?

2 Analytic functions of one matrix

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 a from Ω onto the open unit disk $D = \{z \in \mathbb{C}\ ;\ |z| < 1\}$. From the Osgood-Carathéodory theorem a admits an extension which is a homeomorphism from $\overline{\Omega}$ onto $\overline{D}$. It is convenient to introduce the Blaschke functions

$$b_\zeta(z) = \frac{a(\zeta) - a(z)}{1 - \overline{a(\zeta)}\, a(z)}, \quad \zeta \in \Omega.$$

It is clear that $b_\zeta \in \mathcal{H}^\infty(\Omega) \cap C^0(\overline{\Omega})$, $\|b_\zeta\|_{L^\infty(\Omega)} = 1$ and if $z \in \partial\Omega$, then $|b_\zeta(z)| = 1$. We introduce also the set of finite Blaschke products

$$\mathcal{B}_k(\Omega) = \{f\ ;\ f(z) = e^{i\varphi} \prod_{j=1}^{r} b_{\zeta_j}(z), \quad \varphi \in \mathbb{R},\ \zeta_j \in \Omega,\ j = 1, \dots, r,\ 0 \leq r \leq k\}.$$

(We use the convention $f(z) = e^{i\varphi}$ if $r = 0$, so $\mathcal{B}_0(\Omega)$ corresponds to constant functions of modulus 1). It can be seen that this space is independent of the choice of the isomorphism a between Ω and D. The following theorem is a consequence of the Nevanlinna-Pick theory [8], [9]

Theorem 2. *Let $A \in \mathbb{C}^{d,d}$ be a square matrix with $Sp(A) \subset \Omega$. Then there exists a function $f \in \mathcal{H}^\infty(\Omega)$ such that $\|f(A)\| = \psi_\Omega(A)$ and $\|f\|_{L^\infty(\Omega)} \leq 1$. Furthermore this function has to satisfy $f \in \mathcal{B}_{d-1}(\Omega)$ and $\|f\|_{L^\infty(\Omega)} \leq 1$.*

Proof. Let $\{\lambda_j\}_{j=1,\dots,k}$ the set of the distinct eigenvalues of A and $\{r_j\}$ the corresponding multiplicities. Writting A in Jordan form we can see that $f(A)$ only depends on A and on the values $\{f^{(\ell)}(\lambda_j),\ 0 \leq \ell \leq r_j,\ 1 \leq j \leq k\}$. From the compactness property of $\mathcal{H}^\infty(\Omega)$ we deduce that there exists $f \in \mathcal{H}^\infty(\Omega)$ 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

$$g \in \mathcal{H}^\infty(\Omega), \qquad g^{(\ell)}(\lambda_j) = f^{(\ell)}(\lambda_j), \qquad 0 \leq \ell \leq r_j,\ 1 \leq j \leq k. \tag{4}$$

It is known [10] [12] that there exists $c \in \mathbb{R}^*$ such that $g/c \in \mathcal{B}_{d-1}(\Omega)$; we have $c = \|g\|_{L^\infty(\Omega)} \leq \|f\|_{L^\infty(\Omega)} = 1$ and g is unique. But (4) implies $g(A) = f(A)$, therefore $\psi_\Omega(A) = \|g(A)\| \leq \psi_\Omega(A)/\|g\|_{L^\infty(\Omega)}$. We deduce $\|g\|_{L^\infty(\Omega)} = 1$ and $g = f$. $\square$

Remarks 1. We remark on some properties of the function $\psi_\Omega(A)$.

a) If a denotes a holomorphic isomorphism from Ω onto the unit disk D we have $\psi_\Omega(A) = \psi_D(a(A))$. Indeed $f(A) = g(a(A))$ with $g := f \circ a^{-1}$, 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 H is unitary. That allows to restrict our study to the case of upper triangular matrices without loss of generality.

Lemma 3. *The bound $\psi_\Omega(A)$ depends continuously on A and is (strictly if $\psi_\Omega(A) \neq 1$) decreasing with respect to Ω. Furthermore*

$$\psi(A) = \sup\{\psi_\Omega(A)\ ;\ W(A) \subset \Omega\}. \tag{5}$$

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

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

where $\mathcal{C} \subset \Omega$ 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 functions f bounded by 1.

b) If we have $Sp(A) \subset \Omega \subsetneq \Omega'$, then we have $\psi_{\Omega}(A) > \psi_{\Omega'}(A)$. Indeed let $f \in \mathcal{B}_{d-1}(\Omega')$ such that $\|f(A)\| = \psi_{\Omega'}(A)$ and $\|f\|_{L^\infty(\Omega')} = 1$. We have clearly $\|f(A)\| \le \psi_{\Omega}(A)$ since $\|f\|_{L^\infty(\Omega)} \le 1$, and we cannot have $\|f(A)\| = \psi_{\Omega}(A)$ except if $f \in \mathcal{B}_{d-1}(\Omega) \cap \mathcal{B}_{d-1}(\Omega')$, i.e. if f 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 depends on x. If Ω' is another convex set close to Ω we have

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

That implies continuity with respect to Ω for the Haussdorff distance.

c) Turning now to the proof of (5), we first look at the case when the matrix A is normal. Then we have $\psi(A) = \psi_{\Omega}(A) = 1$, $\forall\ \Omega \supset Sp(A)$ and the result is straightforward. In the other cases the interior of $W(A)$ is not empty and we set $\Omega = \mathrm{int}(W(A))$. If $Sp(A) \subset \Omega$ then we have clearly $\psi(A) \le \psi_{\Omega}(A)$ and from the previous theorem there exists a function $f \in \mathcal{B}_{d-1}(\Omega)$ such that $\psi_{\Omega}(A) = \|f(A)\|$. But $\mathcal{B}_{d-1}(\Omega) \subset C^0(\overline{\Omega})$, thus we can find a sequence of polynomials p_n which uniformly tends to f in Ω which proves that $\psi_{\Omega}(A) \le \psi(A)$. When $Sp(A) \cap \partial\Omega \ne \emptyset$ we can find [6] a unitary matrix U such that $A = U^* \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} U$ with A_2 diagonal and $Sp(A_1) \subset \Omega$. Then it is clear that $\psi(A) = \psi(A_1)$ and for all $\Omega' \supset Sp(A_1)$, $\psi_{\Omega'}(A) = \psi_{\Omega'}(A_1)$. Therefore the result follows from the previous case. □

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

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

THEOREM 4. *If* $T = \begin{pmatrix} \lambda_1 & \gamma \\ 0 & \lambda_2 \end{pmatrix}$, *then we have*

$$\begin{aligned}
\psi_D(T) &= \max\Big(1, \Theta\big(\tfrac{|\gamma|}{|\lambda_1-\lambda_2|}, \tfrac{\sqrt{(1-|\lambda_1|^2)(1-|\lambda_2|^2)}}{|\lambda_1-\lambda_2|}\big)\Big), && \text{if } \lambda_1 \ne \lambda_2,\\
\psi_D(T) &= \max\big(1, \tfrac{|\gamma|}{1-|\lambda_1|^2}\big), && \text{if } \lambda_1 = \lambda_2.
\end{aligned}$$

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

$$\Psi_1(T) = \frac{|\gamma|}{|\lambda_1-\lambda_2|}, \qquad \Psi_2(T) = \frac{\sqrt{(1-|\lambda_1|^2)(1-|\lambda_2|^2)}}{|\lambda_1-\lambda_2|},$$

and we recall that

$$a(T) = \begin{pmatrix} a(\lambda_1) & \gamma\, a[\lambda_1,\lambda_2] \\ 0 & a(\lambda_2) \end{pmatrix}, \quad \text{where } a[\lambda_1,\lambda_2] := \frac{a(\lambda_1) - a(\lambda_2)}{\lambda_1 - \lambda_2}. \tag{6}$$

It is easily verified that if a 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 $a(T)$. Indeed we have just to verify this for automorphisms of the form $a(z) = e^{i\varphi}z$, $\varphi \in \mathbb{R}$ and $a(z) = (x-z)/(1+xz)$, $x \in [0,1)$, since such mappings a generate the automorphism group of D. Remark also that $\psi_{_D}(T)$ does not change if we replace γ by $|\gamma|$ since $\psi_{_D}$ is invariant by a unitary similarity. For all λ_1 and $\lambda_2 \in D$ it is possible to find an automorphism a such that $a(\lambda_1)+a(\lambda_2) = 0$ and $a(\lambda_1) \in (0,1)$. Therefore it is sufficient to prove the theorem in the case where the matrix T is of the form

$$T = \begin{pmatrix} \lambda & 2\,\delta \\ 0 & -\lambda \end{pmatrix}, \qquad \lambda \in (0,1), \quad \delta \geq 0. \tag{7}$$

b) We now consider this case (7). 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

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

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 $\psi_{_D}(T) \leq \|H\|\psi_{_D}(B)\|H^{-1}\| = \|H\|\,\|H^{-1}\|$. But this quantity $\|H\|\,\|H^{-1}\|$ is the largest root of the equation

$$X^2 - \frac{1+\mu^2+\beta^2}{\beta}X + 1 = 0. \tag{8}$$

If we define $X_1 = \sqrt{\delta^2+\lambda^2} + \delta$, $X_2 = 1/X_1$, then we have $X_1\,X_2 = 1$ 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 (8), thus $\|H\|\,\|H^{-1}\| = X_1 = \|T\|$, which implies the theorem. □

Corollary 5. *If* $T = \begin{pmatrix} \lambda_1 & \gamma \\ 0 & \lambda_2 \end{pmatrix}$, *and if a denotes a holomorphic bijection from Ω onto the unit disk D, then we have*

$$\begin{aligned} \psi_{_\Omega}(T) &= \max\big(1, \Theta(\tfrac{|\gamma|}{|\lambda_1-\lambda_2|}, \tfrac{\sqrt{(1-|a(\lambda_1)|^2)(1-|a(\lambda_2)|^2)}}{|a(\lambda_1)-a(\lambda_2)|})\big), && \text{if } \lambda_1 \neq \lambda_2, \\ \psi_{_\Omega}(T) &= \max(1, \tfrac{|\gamma|\,|a'(\lambda_1)|}{1-|a(\lambda_1)|^2}), && \text{if } \lambda_1 = \lambda_2. \end{aligned}$$

Proof. From Remark 1 a) we have $\psi_{_\Omega}(T) = \psi_{_D}(a(T))$. We apply the previous theorem and use Formula (6). □

We turn to the value of the bound $\psi(A)$: we know that $\psi(A) = 1$ if the matrix A is normal. Otherwise it is known that $W(A)$ is an ellipse, and we have

THEOREM 6. *Let A be a non normal 2×2 matrix with entries in $\mathbb{C}$ and let ε be the excentricity of the ellipse $W(A)$. Then $\psi(A) = 2$ if $W(A)$ is a disk (i.e. if $\varepsilon = 0$), otherwise*

$$\psi(A) = 2\exp(-\sum_{n\geq 1} \frac{(-1)^{n+1}}{n} \frac{1}{1+\rho^{4n}}), \quad \text{where } \rho = \frac{1+\sqrt{1-\varepsilon^2}}{\varepsilon}. \tag{9}$$

Proof. We look only at the elliptic case, the disk case will follow by continuity. In order to prove (9) we first remark that ψ_A does not change if we replace A by $\alpha A + \beta$, $\alpha \neq 0$, α, $\beta \in \mathbb{C}$, or by a unitarily similar matrix. Therefore we can assume that $A = \begin{pmatrix} 1 & \gamma \\ 0 & -1 \end{pmatrix}$ and furthermore that $\gamma > 0$. Then $W(A)$ is the ellipse of foci 1, -1 and minor axis γ (see [3] for instance). If ε denotes its excentricity, we have $\gamma = \rho - 1/\rho$ and the major axis is $\rho + 1/\rho$; furthermore (see [4]) the function

$$a(z) = \frac{2z}{\rho} \exp\big(-\sum_{n\geq 1} \frac{(-1)^{n+1}}{n} \frac{t_{2n}(z)}{1+\rho^{4n}}\big),$$

(where t_n denotes the n^{th} Chebyshev polynomial) is the Riemann conforming function mapping $W(A)$ onto $\overline{D}$. We deduce from Corollary 5 that, with $\Omega = \mathrm{int}(W(A))$,

$$\psi(A) = \psi_\Omega(A) = \max\big(1, \Theta\big(\frac{\gamma}{2}, \frac{1-|a(1)|^2}{2\,|a(1)|}\big)\big) = \rho\, a(1),$$

which gives (9). □

Theorem 1 is a direct consequence of this result since from (9) we clearly have $\psi(A) \leq 2$.

3 Analytic functions of a family of matrices

This section is concerned with the study of the bound $C_\Omega(d)$. For the two dimensional case it follows from the last theorem that

$$C_\Omega(2) = \sup\{\psi_\Omega(A)\,;A \in \mathcal{M}_2(\Omega)\} \leq \sup\{\psi(A)\,;A \in \mathcal{M}_2(\Omega)\} \leq 2. \tag{10}$$

We do not know if the same result holds in higher dimension.

REMARK 2. In [5] the authors give the majoration $C_D(d) \leq 3$ for the disk and in the general case

$$C_\Omega(d) \leq 3 + \Big(\frac{2\pi\delta^2}{|\Omega|}\Big)^3,$$

where δ is the diameter of the bounded convex set Ω and $|\Omega|$ its area. This estimate (which is very likely pessimistic) implies the interesting property

$\forall\, f \in \mathcal{H}^\infty(\Omega)\cap C^0(\overline{\Omega})$ the map $A \mapsto f(A)$ is uniformly continuous from $\mathcal{M}_d(\Omega)$ into $\mathbb{C}^{d,d}$,

indeed this property is true when f is a polynomial and the polynomials are dense in $\mathcal{H}^\infty(\Omega)\cap C^0(\overline{\Omega})$.

The following theorems concern the realization of the bound $C_\Omega(\Omega)$.

THEOREM 7. *Let Ω be a bounded convex subset of the complex plane with an analytic boundary. Then there exists a matrix $A \in \mathcal{M}_d(\Omega)$ and a function $f \in \mathcal{B}_{d-1}(\Omega)$ such that*

$$C_\Omega(d) = \|f(A)\|.$$

Proof. The case $d = 1$ is trivial, $C_\Omega(1) = 1$ and this constant is realized with the constant function $f = 1$. We clearly have $C_\Omega(d) \geq C_\Omega(d-1)$. The case $C_\Omega(d) = C_\Omega(d-1)$ can be treated by induction. Therefore we only look from now on at the case

$$C_\Omega(d) > C_\Omega(d-1). \tag{11}$$

From Theorem 2 there exists a sequence of $f_n \in \mathcal{B}_{d-1}(\Omega)$ and $A_n \in \mathbb{C}^{d,d}$ satisfying

$$W(A_n) \subset \overline{\Omega},\ Sp(A_n) \subset \Omega,\ \|f_n\|_{L^\infty(\Omega)} = 1, \quad \text{and} \lim_{n\to\infty} \|f_n(A_n)\| = C_\Omega(d).$$

We write $f_n(z) = \prod_{j=1}^{r_n} b_{\zeta_j^n}(z)$. The sequence of $\{\zeta_j^n \in \Omega\}$ is clearly bounded. The sequence $\{A_n\}$ is also bounded since $\|A_n\| \leq 2\sup\{|z|\,;\ z \in W(A_n)\}$. Therefore (after extraction of a subsequence) we can assume that $A_n \to A$, $r_n = r$, $\zeta_j^n \to \zeta_j$. We set $f(z) = \prod_{j=1}^{r} b_{\zeta_j}(z)$; we have $f \in \mathcal{B}_{d-1}(\Omega)$. Indeed $\zeta_j \in \overline{\Omega}$, and if $\zeta_j \in \partial\Omega$ then $b_{\zeta_j}(z) = a(\zeta_j)$ is constant in Ω. We also have $Sp(A) \subset \overline{\Omega}$ and $W(A) \subset \overline{\Omega}$.

We first consider the case where $Sp(A) \subset \Omega$, i.e. $A \in \mathcal{M}_d(\Omega)$. From Remark 2 we have $f(A_n) \to f(A)$ and therefore

$$\begin{aligned}\|f_n(A_n) - f(A)\| &\leq \|f(A_n) - f(A)\| + \|f_n(A_n) - f(A_n)\| \\ &\leq \|f(A_n) - f(A)\| + C_\Omega(d)\|f_n - f\|_{L^\infty(\Omega)} \to 0,\end{aligned}$$

and consequently $\|f(A)\| = C_\Omega(d)$. Thus the theorem holds if we prove that the opposite case $Sp(A) \cap \partial\Omega \neq \emptyset$ is impossible.

If $Sp(A) \cap \partial\Omega \neq \emptyset$, let $\lambda \in \partial\Omega$ an eigenvalue of A and k its multiplicity. We can assume that the matrices A_n (and thus A) were chosen upper triangular and also that $a_{11} = \cdots = a_{kk} = \lambda$. Then we remark that $\forall i = 1, \ldots, k$, $\forall j > i$, we have $a_{ij} = 0$. Indeed (using the canonical basis $\{e_j\}$ in $\mathbb{C}^d$) the condition $W(A) \subset \overline{\Omega}$ implies

$$\forall \mu \in \mathbb{C}, \qquad \lambda + \mu a_{ij} + |\mu|^2 a_{jj} = (e_i + \mu e_j)^* A (e_i + \mu e_j) \in (1 + |\mu|^2)\overline{\Omega},$$

which is only compatible with $\lambda \in \partial\Omega$ in the case $a_{ij} = 0$. Now we write the matrices in block form

$$A_n = \begin{pmatrix} T_n & C_n \\ 0 & S_n \end{pmatrix}, \quad A = \begin{pmatrix} \lambda I_k & 0 \\ 0 & S \end{pmatrix}.$$

For n large enough $S_p(T_n) \cap S_p(S_n) = \emptyset$, therefore we can define U_n as the unique solution of

$$U_n S_n - T_n U_n = C_n,$$

which we can also write as

$$U_n = (C_n + (T_n - \lambda I_k)U_n)(S_n - \lambda I_{d-k})^{-1}.$$

Since $C_n \to 0$ and $T_n - \lambda I_k \to 0$ we deduce $U_n \to 0$. We remark now that

$$A_n = \begin{pmatrix} I & -U_n \\ 0 & I \end{pmatrix} \begin{pmatrix} T_n & 0 \\ 0 & S_n \end{pmatrix} \begin{pmatrix} I & U_n \\ 0 & I \end{pmatrix},$$

thus

$$f_n(A_n) = \begin{pmatrix} I & -U_n \\ 0 & I \end{pmatrix} \begin{pmatrix} f_n(T_n) & 0 \\ 0 & f_n(S_n) \end{pmatrix} \begin{pmatrix} I & U_n \\ 0 & I \end{pmatrix}.$$

But we have $\|f_n(T_n)\| \le C_\Omega(k)$ and $\|f_n(S_n)\| \le C_\Omega(d-k)$. If $k < d$ this implies $\|f_n(A_n)\| \le (1 + \|U_n\|)C_\Omega(d-1)(1 + \|U_n\|)$. Taking the limit as $n \to \infty$ we deduce $C_\Omega(d) \le C_\Omega(d-1)$ which is in contradiction with (11). Therefore the only possible case is $k = d$ i.e. $A = \lambda I_d$.

It remains to show that this case $A = \lambda I_d$ is impossible. Note that $a \in C^0(\overline{\Omega})$. As in Remark 2 we deduce from $A_n \to \lambda I_d$ that $a(A_n) \to a(\lambda)\, I_d$, therefore if $\zeta_j \ne \lambda$ it follows from $\zeta_j^n \to \zeta_j$ that

$$b_{\zeta_j^n}(A_n) = (a(\zeta_j^n) - a(A_n))(1 - \overline{a(\zeta_j^n)}\, a(A_n))^{-1} \to b_{\zeta_j}(\lambda)\, I_d.$$

If $\zeta_j = \lambda$, we use the smoothness assumption on the boundary. The function a admits a holomorphic extension into a neighborhood of λ. Let C be a circle of center λ included in this neighborhood. We deduce from the uniform convergence of $b_{\zeta_j^n}$ towards $b_\lambda = a(\lambda) = b_\lambda(\lambda)$ on D and from the Cauchy representation

$$b_{\zeta_j^n}(A_n) = \frac{1}{2\pi i} \int_C (z - A_n)^{-1}\, b_{\zeta_j^n}(z)\, dz,$$

that $b_{\zeta_j^n}(A_n) \to b_\lambda(\lambda)\, I_d$. Therefore we have $\lim_{n\to\infty} f_n(A_n) = f(\lambda)\, I_d$, and consequently $C_\Omega(d) = 1$ which is also in contradiction with (11) since $C_\Omega(d-1) \ge 1$. $\square$

THEOREM 8. *Let Ω be a strip or a convex sector of the complex plane. Then there exists a matrix $A \in \mathcal{M}_d(\Omega)$ and a function $f \in \mathcal{B}_{d-1}(\Omega)$ with $\|f\|_{L^\infty(\Omega)} = 1$ such that*

$$C_\Omega(d) = \|f(A)\|.$$

Proof. We first consider the strip case. Without loss of generality we can assume that Ω is the strip

$$\Omega = B = \{z \in \mathbb{C};\, |\operatorname{Im} z| \le 1\}, \qquad \text{and we use } a(z) = \tanh \frac{\pi}{4} z.$$

Note that (after a slight modification which is the multiplication by a unitary constant) the Blaschke functions are now $b_\zeta(z) = \frac{\mathrm{sh}(\frac{\pi}{4}(\zeta - z))}{\mathrm{ch}(\frac{\pi}{4}(\overline{\zeta} - z))}$. As in the bounded case we can restrict our study to the case

$$C_\Omega(d) > C_\Omega(d-1).$$

Then we consider a sequence of $f_n \in \mathcal{B}_{d-1}(B)$ and $A_n \in \mathbb{C}^{d,d}$ satisfying

$$W(A_n) \subset \overline{B},\ Sp(A_n) \subset B,\ \|f_n\|_{L^\infty(\Omega)} = 1, \quad \text{and } \lim_{n\to\infty} \|f_n(A_n)\| = C_\Omega(d). \tag{12}$$

We write $f_n(z) = \prod_{j=1}^{r_n} b_{\zeta_j^n}(z)$. We can assume A_n upper triangular, with the eigenvalues $\lambda_j^n = a_{jj}^n$ ordered such that $\operatorname{Re} \lambda_j^n$ increase with j. We can also suppose $\operatorname{trace}(A_n) = 0$ (indeed replacing (A_n, f_n) by $(A_n - x_n, f_n(. + x_n)))$ for $x_n \in \mathbb{R}$ does not modify (12)).

We first prove that the sequence A_n is bounded. Otherwise (eventually after extraction of a subsequence) there will exist an index k such that $\mathrm{Re}(\lambda_{k+1}^n - \lambda_k^n) \to \infty$. Then we write in block form

$$A_n = \begin{pmatrix} T_n & C_n \\ 0 & S_n \end{pmatrix}, \qquad \text{with } T_n \in \mathbb{C}^{k,k}.$$

Since $Sp(S_n) \cap Sp(T_n) = \emptyset$ we can define U_n as the unique solution of

$$U_n S_n - T_n U_n = C_n,$$

which we can also write ($\bullet$ denotes the componentwise product)

$$U_n = H_n \bullet \big(C_n + (T_n - \mathrm{diag}(T_n))\, U_n - U_n\, (S_n - \mathrm{diag}(S_n))\big),$$

where $\qquad H_n = (h_{ij}^n) \in \mathbb{C}^{k,d-k}$ is defined by $h_{ij}^n = \frac{1}{\lambda_i^n - \lambda_{j+k}^n}$.

We have $W(A_n) \subset \overline{B}$, i.e. $\|A_n - A_n^*\| \le 2$. Since A_n is triangular, this implies that $A_n - \mathrm{diag}(A_n)$ is bounded. Therefore C_n, $T_n - \mathrm{diag}(T_n)$ and $S_n - \mathrm{diag}(S_n)$ are bounded. Joint with $H_n \to 0$, this implies $U_n \to 0$. We remark now that

$$A_n = \begin{pmatrix} I & -U_n \\ 0 & I \end{pmatrix} \begin{pmatrix} T_n & 0 \\ 0 & S_n \end{pmatrix} \begin{pmatrix} I & U_n \\ 0 & I \end{pmatrix},$$

thus

$$f_n(A_n) = \begin{pmatrix} I & -U_n \\ 0 & I \end{pmatrix} \begin{pmatrix} f_n(T_n) & 0 \\ 0 & f_n(S_n) \end{pmatrix} \begin{pmatrix} I & U_n \\ 0 & I \end{pmatrix}.$$

This implies $\|f_n(A_n)\| \le (1 + \|U_n\|)^2 C_\Omega(d-1)$, and by taking the limit at ∞ gives $C_\Omega(d) \le C_\Omega(d-1)$ which is in contradiction with (11). Therefore the sequence A_n is bounded.

We turn now to the sequences $\{\zeta_j^n\}$. After extraction of a subsequence we can assume that $r_n = r$ and that for each j one of these three possibilities occurs : "the sequence $\{\zeta_j^n\}$ is bounded", " $\lim \mathrm{Re}\, \zeta_j^n = +\infty$"," $\lim \mathrm{Re}\, \zeta_j^n = -\infty$". We can also assume that $\limsup \|b_{\zeta_j^n}(A_n)\| > 1$ (otherwise we could replace f_n by $f_n / b_{\zeta_j^n}$ and (12) would be still satisfied). We consider now the cases " $\lim \mathrm{Re}\, \zeta_j^n = \varepsilon\, \infty$" ($\varepsilon = \pm 1$). Then we have $b_{\zeta_j^n} \to \varepsilon$ uniformly on each compact subset of $\mathbb{C}$. Consequently $\lim \|b_{\zeta_j^n}(A_n)\| = 1$, possibility which has been excluded. Therefore we have proved that the family $\{\zeta_j^n\}$ is bounded. The remaining part of the proof is the same as for Theorem 7.

We turn now to the proof of Theorem 8 in the case of a convex sector Ω. We can assume that $\Omega = S_a := \{z \in \mathbb{C};\, z \ne 0 \text{ and } |\arg z| < \alpha\}$, for some $0 < \alpha \le \frac{\pi}{2}$, and that $C_\Omega(d) > C_\Omega(d-1)$. We consider a sequence of functions $f_n \in \mathcal{B}_{d-1}(S_\alpha)$ and upper triangular matrices $A_n \in \mathbb{C}^{d,d}$ satisfying

$$W(A_n) \subset \overline{S}_\alpha,\ Sp(A_n) \subset S_\alpha,\ \|f_n\|_{L^\infty(S_\alpha)} = 1,\ \lim_{n\to\infty} \|f_n(A_n)\| = C_{S_\alpha}(d). \tag{13}$$

We write $f_n(z) = \prod_{j=1}^{r_n} b_{\zeta_j^n}(z)$. We also assume that $|\det(A_n)| = 1$ (this is possible since replacing (A_n, f_n) by $(x_n A_n, f_n(./x_n))$ with $x_n > 0$ does not affect (13) .

We introduce now $L_n = \frac{1}{\alpha}\log(A_n)$, $g_n(\zeta) = f_n(e^{\alpha z})$. Then we have $g_n \in \mathcal{B}_{d-1}(S_B)$, $L_n \in \mathbb{C}^{d,d}$ satisfying

$$W(L_n) \subset \overline{B},\ Sp(L_n) \subset B,\ \|g_n\|_{L^\infty(B)} = 1, \quad \text{and } \lim_{n\to\infty} \|g_n(L_n)\| = C_{S_\alpha}(d). \tag{14}$$

(the proof of $W(L_n) \subset \overline{B}$ will be given in Theorem 9). The argumentation used in the strip case allows to show that the sequence $\{L_n\}$ is bounded, consequently the sequences $\{A_n\}$ and $\{(A_n)^{-1}\}$ are bounded. Similarly we obtain the boundedness of the sequences $\{\zeta_j^n\}$ and $\{(\zeta_j^n)^{-1}\}$. We conclude as previously. □

THEOREM 9. *Let S_α be a sector with angle $0 \leq 2\alpha \leq \pi$ and B be a strip of the complex plane. Then the constant $C_{S_\alpha}(d)$ is a decreasing function of α. We have $C_{S_\alpha}(d) \leq C_B(d)$ and $C_B(d) = \lim_{\alpha\to 0} C_{S_\alpha}(d)$.*

Proof. a) We first prove that $W(A) \subset S_\alpha$ implies $W(A^s) \subset S_{s\alpha}$ for all $0 < s < 1$. For this we introduce the two half planes $\Pi_{\alpha+}$ and $\Pi_{\alpha-}$

$$\Pi_{\alpha+} = \{z \in \mathbb{C}; \mathrm{Re}\, e^{i(\pi/2-\alpha)}\, z > 0\},\ \ \Pi_{\alpha-} = \{z \in \mathbb{C}; \mathrm{Re}\, e^{i(-\pi/2+\alpha)}\, z > 0\},$$

so that $S_\alpha = \Pi_{\alpha+} \cap \Pi_{\alpha-}$. It is clear that the function $a(z) = z^s$ (we choose the branch such that $a(1) = 1$) maps the half plane $\Pi_{\alpha\pm}$ into $\Pi_{s\alpha\pm}$. Since $W(A) \subset \Pi_{\alpha\pm}$ we have (still a von Neumann result) $W(a(A)) \subset a(\Pi_{\alpha\pm})$, i.e. $W(A^s) \subset S_{s\alpha}$.

b) A first consequence is that $C_{S_\alpha}(d)$ decreases with α. Indeed we choose a $d\times d$ matrix A such that $W(A) \subset \overline{S}_\alpha$ and $\psi_{S_\alpha}(A) = C_{S_\alpha}(d)$. We have, for $0 < s < 1$, $\psi_{S_\alpha}(A) = \psi_{S_{s\alpha}}(A^s) \leq C_{S_{s\alpha}}(d)$ since $W(A^s) \subset \overline{S}_{s\alpha}$.

c) With the same choice of matrix we set $L = \frac{1}{\alpha}\log(A)$. It is clear that $a(z) = \frac{1}{\alpha}\log z$ is an isomorphism from S_α onto B. Thus we have $\psi_{S_\alpha}(A) = \psi_B(L)$. We will get $\psi_B(L) \leq C_B(d)$ i.e. $C_{S_\alpha}(d) \leq C_B(d)$ if we prove that $W(L) \subset \overline{B}$.

For this we remark that $L = \lim_{n\to\infty} \frac{n}{\alpha}(A^{1/n} - I)$. Therefore if $u \in \mathbb{C}^d$ satisfies $\|u\| = 1$ we have (since $W(A^{1/n}) \subset S_{\alpha/n}$)

$$|\,\mathrm{Im}\,\langle Lu, u\rangle| = \lim_{n\to\infty} \frac{n}{\alpha}\ |\,\mathrm{Im}\,\langle A^{1/n}u, u\rangle| \leq \lim_{n\to\infty} \frac{n}{\alpha} \tan(\frac{\alpha}{n})\ \mathrm{Re}\,\langle A^{1/n}u, u\rangle = 1,$$

which gives $W(L) \subset \overline{B}$.

d) We know that $C_B(d) \geq \lim_{\alpha\to 0} C_{S_\alpha}(d)$, and it remains to prove that $C_B(d) \leq \lim_{\alpha\to 0} C_{S_\alpha}(d)$. For this it is sufficient to show that for any $d\times d$ matrix L with $W(L) \subset B$ there exists a positive α such that $\psi_B(L) \leq C_{S_\alpha}(d)$. Let $L = M + iN$ be such a matrix with M and N self-adjoint ; since $W(L) \subset B$ we have $\|N\| < 1$. We set $A = \exp(\alpha L)$, then $\psi_B(L) = \psi_{S_\alpha}(A)$. Furthermore if $u \in \mathbb{C}^d$ satisfies $\|u\| = 1$ we have

$$\langle Au, u\rangle = 1 + \alpha\,\langle Mu, u\rangle + i\alpha\,\langle Nu, u\rangle + O(\alpha^2),\ \text{ thus } \tan(\arg(\langle Au, u\rangle)) = \alpha\,\langle Nu, u\rangle + O(\alpha^2),$$

and consequently $|\tan(\arg(\langle Au, u\rangle))| \leq \alpha\,(\|N\| + O(\alpha))$ (the term $O(\alpha)$ tends to 0 uniformly in u). Therefore we can choose α such that $\alpha\,(\|N\| + O(\alpha)) \leq \tan\alpha$, and then we have $W(A) \subset \overline{S}_\alpha$ and consequently $\psi_B(L) = \psi_{S_\alpha}(A) \leq C_{S_\alpha}(d)$. □

THEOREM 10. *In the disk case, for $d \geq 2$ we have $C_D(d) = 2$.*

Proof. We sketch the proof proposed to us by [2]. We start from a result of [1] which asserts that, if $A \in \mathcal{M}_d(D)$, then there exist two matrices B and C satisfying B is hermitian, $\|C\| \leq 1$, such that $A = 2\cos(B)C\sin(B)$. We consider the continuous function $g(x) = \max(1, 2|\cos x|)$ and define (using the continuous functional calculus for self-adjoint operators) the matrices $H = g(B)$ and $T = H^{-1}AH$. We clearly have

$$\|H\| \leq 2, \quad \|H^{-1}\| \leq 1, \quad \|\sin(B)H\| \leq 1, \quad 2\,\|H^{-1}\cos(B)\| \leq 1, \quad \text{thus } \|T\| \leq 1.$$

If $f \in \mathcal{H}^\infty(D)$ we have $f(A) = Hf(T)H^{-1}$, therefore $\|f(A)\| \leq \|H\|\|f(T)\|\|H^{-1}\| \leq 2\|f\|_{L^\infty(D)}$. With the choice $f(z) = z$ and $A = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$ we see that the bound 2 is attained.

□

4 The two dimensional case

In this section we consider 2×2 matrices.

Theorem 11. *For any open convex Ω ($\neq \emptyset$ and $\neq \mathbb{C}$) of the complex plane we have $C_\Omega(2) \leq 2$. Furthermore the equality $C_\Omega(2) = 2$ implies that Ω is a disk.*

Proof. The inequality has been seen previously (10). If there exists a matrix $A \in \mathcal{M}_2(\Omega)$ such that $C_\Omega(2) = \psi_\Omega(A)$ (which is the case if the boundary of Ω is analytic, cf. Theorem 7) then

$$C_\Omega(2) = \psi_\Omega(A) \leq \psi(A) \leq 2.$$

But the first inequality is strict if $\overline{\Omega} \neq W(A)$, and from (9) the second is strict if Ω is not a disk. Therefore in the analytic case $C_\Omega(2) = 2$ implies that Ω is a disk.

We turn now to the case when $C_\Omega(2) = 2$ and the constant $C_\Omega(2)$ is not attained. Arguing as in the proof of Theorem 7 we can find a sequence of upper triangular matrices $A_n \in \mathcal{M}_2(\Omega)$ and of functions f_n with the following properties

$$f_n(z) = \tfrac{a(\zeta_n)-a(z)}{1-\overline{a(\zeta_n)}a(z)}, \quad \|f_n(A_n)\| \to 2, \quad \zeta_n \to \zeta, \quad \text{and } A_n \to \zeta\, Id.$$

We set

$$A_n = \begin{pmatrix} \lambda_1^n & 2\,\varepsilon_n \\ 0 & \lambda_2^n \end{pmatrix}, \qquad \lambda_n = \frac{\lambda_1^n + \lambda_2^n}{2}, \qquad B_n = \begin{pmatrix} \lambda_n & 2\,\varepsilon_n \\ 0 & \lambda_n \end{pmatrix}.$$

We have $\psi_{A_n} \to 2$ and we deduce from (9) that the excentricity of $W(A_n)$ tends to 0, thus $\frac{\lambda_1^n - \lambda_2^n}{\varepsilon_n} \to 0$. Using the Cauchy representation

$$f_n(A_n) - f_n(B_n) = \frac{1}{2\pi i} \int_{\partial W(B_n)} f_n(z)\,((z-A_n)^{-1} - (z-B_n)^{-1})\,dz,$$

we get $f_n(A_n) - f_n(B_n) \to 0$. Since the disk $W(B_n) = \{z\,;\, |z-\lambda_n| \leq \varepsilon_n\}$ satisfies $W(B_n) \subset W(A_n) \subset \overline{\Omega}$ and since Ω is convex it is possible (at least for n large enough) to find a point $\mu_n \in \partial W(B_n)$ such that the disk $\Delta_n = \{z\,;\, |z-\mu_n| < \varepsilon_n\}$ satisfies $\Delta_n \subset \Omega$. We consider now the set $W = \{z\,;\, z = a + \xi,\, 0 < a < 1,\, |\xi| < 1\}$ which is the convex hull of the union of the unit disk D and the disk $\Delta = 1 + D$, and we define

$$g_n(z) := f_n(\lambda_n + z(\mu_n - \lambda_n)), \qquad B := \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}.$$

We clearly have $\|g_n(B)\| = \|f_n(B_n)\| \to 2$ and $\|g_n\|_{L^\infty(W)} \leq 1$. We can extract a subsequence such that $g_n \to g \in \mathcal{H}^\infty(W)$ (in C^∞ of any compact subset of W). Then we have $\|g(B)\| = 2$ and $\|g\|_{L^\infty(W)} \leq 1$. From $\psi_B = 2$ and Theorem 2 we deduce $g \in \mathcal{B}_1(D)$, i.e. we can write $g(z) = e^{i\varphi}\,(z-\mu)(1-\overline{\mu}z)^{-1}$ with $\mu \in D$. But then $2 = \|g(B)\| = 1 + \sqrt{1+|\mu|^2}$ implies $\mu = 0$, i.e. $g(z) = e^{i\varphi}\,z$, which is in contradiction with $\|g\|_{L^\infty(W)} \leq 1$. □

We now give explicit formulae for the sector and for the strip in 2 dimensions.

THEOREM 12. *In the 2-dimensional case we have the following formulae*

$$\begin{array}{rcll} C_{S_\alpha}(2) & = & \sup_{y>0}\ \tau_\alpha(y), & \textit{if } 0 < \alpha < \frac{\pi}{4}, \\ C_{S_\alpha}(2) & = & \sqrt{2}, & \textit{if } \alpha = \frac{\pi}{4}, \\ C_{S_\alpha}(2) & = & \sup_{c\in(0,1)}\ \sigma_\alpha(c), & \textit{if } \frac{\pi}{4} < \alpha \leq \frac{\pi}{2}, \end{array}$$

where

$$\tau_\alpha(y) := \tanh \tfrac{\pi}{4} y\ \frac{\sin\alpha + \sqrt{\sin^2\alpha + \mathrm{sh}^2\,\alpha y}}{\mathrm{sh}\,\alpha y},$$

and

$$\sigma_\alpha(c) := \tan \tfrac{\pi}{4} c\ \frac{\sin\alpha + \sqrt{\sin^2\alpha - \sin^2\alpha c}}{\sin \alpha c}.$$

Proof. In this proof we will use the generic upper triangular matrix T and the isomorphism a from S_α onto D defined by

$$T = \begin{pmatrix} \lambda_1 & \gamma \\ 0 & \lambda_2 \end{pmatrix}, \quad \text{and} \quad a = \frac{z^{\pi/2\alpha} - \lambda_2^{\pi/2\alpha}}{z^{\pi/2\alpha} - \overline{\lambda}_2^{\pi/2\alpha}}.$$

By an argument of continuity it is sufficient to only look at the case $\lambda_1 \neq \lambda_2$. Since $a(\lambda_2) = 0$, it follows from Corollary 5 that

$$\psi_{S_\alpha}(T) = \max\ \big(1, \Theta(\frac{|\gamma|}{|\lambda_1 - \lambda_2|}, \frac{\sqrt{1-|a(\lambda_1)|^2}}{|a(\lambda_1)|})\big). \tag{15}$$

Note that $W(T)$ is an ellipse (cf. [6] for instance) with foci λ_1 and λ_2 and minor axis $|\gamma|$. We remark that it is sufficient to consider the matrices T such that $W(T)$ is tangent to the two edges of S_α, indeed otherwise we could find another sector $S = \zeta + S_\alpha$ such that $W(T) \subset \overline{S} \subsetneq \overline{S}_\alpha$ but then, from Lemma 3, $\psi_{S_\alpha}(T) < \psi_S(T) = \psi_{S_\alpha}(T-\zeta)$. Therefore

$$C_{S_\alpha}(2) = \max\{\psi_{S_\alpha}(T)\,;T \in \mathcal{M}_d,\ W(T) \text{ bitangent to } \partial S_\alpha\}. \tag{16}$$

From the second Poncelet theorem the two edges of S_α (tangent to $W(T)$) and the two focus lines $O\lambda_1$, $O\lambda_2$ have the same bisectrix, therefore $\arg\lambda_1 + \arg\lambda_2 = 0$. Thus we can set

$$\lambda_1 = \rho\, e^{\alpha(y+i\theta)}, \quad \lambda_2 = \rho\, e^{-\alpha(y+i\theta)}, \quad \text{with } y \in \mathbb{R},\ |\theta| < 1.$$

The edge $\{r\,e^{i\alpha}\,; r > 0\}$ being tangent to $W(T)$ we have

$$\begin{aligned} 0 &= \max_{\|u\|=1} \mathrm{Im}\ e^{-i\alpha}\,\langle Tu, u\rangle \\ &= \max_{\|u\|=1} (-\rho\, e^{\alpha y}|u_1|^2 \sin\alpha(1-\theta) - \rho\, e^{-\alpha y}|u_2|^2 \sin\alpha(1+\theta) + \mathrm{Im}\ \gamma\, e^{-i\alpha}\,\bar{u}_1 u_2), \\ &= |\gamma|/2 - \rho\sqrt{\sin\alpha(1-\theta)\ \sin\alpha(1+\theta)}. \end{aligned}$$

Thus $|\gamma| = 2\,\rho\,\sqrt{\cos^2\alpha\theta - \cos^2\alpha}$, furthermore $|\lambda_1 - \lambda_2| = 2\,\rho\,|\operatorname{sh}\alpha(y+i\theta)| = 2\,\rho\,\sqrt{\operatorname{ch}^2\alpha y - \cos^2\alpha\theta}$, this gives

$$\frac{|\gamma|}{|\lambda_1 - \lambda_2|} = \frac{\sqrt{\cos^2\alpha\theta - \cos^2\alpha}}{\sqrt{\operatorname{ch}^2\alpha y - \cos^2\alpha\theta}}.$$

Similarly we get

$$\frac{\sqrt{1-|a(\lambda_1)|^2}}{|a(\lambda_1)|} = \frac{2\sqrt{\operatorname{Re}\lambda_1^{\pi/2\alpha}\ \operatorname{Re}\lambda_2^{\pi/2\alpha}}}{|\lambda_1^{\pi/2\alpha} - \lambda_2^{\pi/2\alpha}|} = \frac{\cos\frac{\pi\theta}{2}}{\sqrt{\operatorname{ch}^2\frac{\pi y}{2} - \cos^2\frac{\pi\theta}{2}}}.$$

Putting this into (15), we deduce if $\psi_{S_\alpha}(T) > 1$

$$\begin{aligned}
\psi_{S_\alpha}(T) &= \frac{\sqrt{\operatorname{ch}^2\alpha y - \cos^2\alpha} + \sqrt{\cos^2\alpha\theta - \cos^2\alpha}}{\sqrt{\operatorname{ch}^2\alpha y - \cos^2\alpha\theta}}\ \frac{\sqrt{\operatorname{ch}^2\frac{\pi y}{2} - \cos^2\frac{\pi\theta}{2}}}{\operatorname{ch}\frac{\pi y}{2} + \cos\frac{\pi\theta}{2}}, \\
&= \left(\frac{\sqrt{\operatorname{ch}^2\alpha y - \cos^2\alpha} + \sqrt{\cos^2\alpha\theta - \cos^2\alpha}}{\sqrt{\operatorname{ch}^2\alpha y - \cos^2\alpha} - \sqrt{\cos^2\alpha\theta - \cos^2\alpha}}\right)^{1/2} \left(\frac{\operatorname{ch}\frac{\pi y}{2} - \cos\frac{\pi\theta}{2}}{\operatorname{ch}\frac{\pi y}{2} + \cos\frac{\pi\theta}{2}}\right)^{1/2},
\end{aligned}$$

and coming back to (16) we get

$$C_{S_\alpha}(2) = \max\{(\phi_\alpha(y,\theta))^{1/2}; y \geq 0, \theta \in [0,1]\}, \quad \text{with } \phi_\alpha(y,\theta) := (\psi_{S_\alpha}(T))^2. \tag{17}$$

More explicitly

$$\begin{aligned}
\phi_\alpha(0,0) &= \frac{\pi\sin\alpha}{2\alpha}, \\
\phi_\alpha(y,\theta) &= \frac{\sqrt{\operatorname{ch}^2\alpha y - \cos^2\alpha} + \sqrt{\cos^2\alpha\theta - \cos^2\alpha}}{\sqrt{\operatorname{ch}^2\alpha y - \cos^2\alpha} - \sqrt{\cos^2\alpha\theta - \cos^2\alpha}}\ \frac{\operatorname{ch}\frac{\pi y}{2} - \cos\frac{\pi\theta}{2}}{\operatorname{ch}\frac{\pi y}{2} + \cos\frac{\pi\theta}{2}}.
\end{aligned}$$

In the case $\alpha = \frac{\pi}{4}$ simple calculations give

$$\phi_{\pi/4}(y,\theta) := \frac{(\sqrt{\operatorname{ch}\frac{\pi}{2}y} + \sqrt{\cos\frac{\pi}{2}\theta})^2}{\operatorname{ch}\frac{\pi}{2}y + \cos\frac{\pi}{2}\theta},$$

and using the Cauchy-Schwarz inequality we deduce $C_{S_{\pi/4}}(2) = \sqrt{2} = \sqrt{\phi_{\pi/4}(0,0)}$.

We turn to the other values of α. We remark that the function ϕ_α is C^∞ except maybe in the point $(y,\theta) = (0,0)$. We know from Theorem 8 that the constant $C_{S_\alpha}(2)$ is realized for a matrix T. Therefore there is a point (y,θ) such that either $(y,\theta) = (0,0)$, or $\nabla\phi_\alpha(y,\theta) = 0$. This last relation may be written

$$\begin{aligned}
\frac{\alpha\operatorname{sh}\alpha y\ \operatorname{ch}\alpha y}{\sqrt{\operatorname{ch}^2\alpha y - \cos^2\alpha}}\ \frac{2\sqrt{\cos^2\alpha\theta - \cos^2\alpha}}{\operatorname{ch}^2\alpha y - \cos^2\alpha\theta} &= \frac{\pi}{2}\operatorname{sh}\frac{\pi y}{2}\ \frac{2\cos\frac{\pi\theta}{2}}{\operatorname{ch}^2\frac{\pi y}{2} - \cos^2\frac{\pi\theta}{2}}, \\
\frac{\alpha\sin\alpha\theta\ \cos\alpha\theta}{\sqrt{\cos^2\alpha\theta - \cos^2\alpha}}\ \frac{2\sqrt{\operatorname{ch}^2\alpha y - \cos^2\alpha}}{\operatorname{ch}^2\alpha y - \cos^2\alpha\theta} &= \frac{\pi}{2}\sin\frac{\pi\theta}{2}\ \frac{2\operatorname{ch}\frac{\pi y}{2}}{\operatorname{ch}^2\frac{\pi y}{2} - \cos^2\frac{\pi\theta}{2}}.
\end{aligned}$$

If $y \neq 0$ and $\theta \neq 0$ this implies

$$\frac{\operatorname{sh}2\alpha y}{\sin 2\alpha\theta}\ \frac{\cos^2\alpha\theta - \cos^2\alpha}{\operatorname{ch}^2\alpha y - \cos^2\alpha} = \frac{\tanh\frac{\pi y}{2}}{\tan\frac{\pi\theta}{2}},$$

or equivalently $\mu_\alpha(y) = \nu_\alpha(\theta)$ with

$$\mu_\alpha(y) := \tanh\frac{\pi y}{2}\ \frac{\operatorname{ch}2\alpha y - \cos 2\alpha}{\operatorname{sh}2\alpha y}, \qquad \nu_\alpha(\theta) = \tan\frac{\pi\theta}{2}\ \frac{\cos 2\alpha\theta - \cos 2\alpha}{\sin 2\alpha\theta}.$$

But this is impossible since for $y > 0$ and $\theta \in (0,1)$ we have (see Lemma 13 below)

$$\nu_\alpha(\theta) < \tfrac{\pi \sin^2\alpha}{2\alpha} < \mu_\alpha(y), \qquad \text{if } 0 < \alpha < \tfrac{\pi}{4},$$

$$\nu_\alpha(\theta) > \tfrac{\pi \sin^2\alpha}{2\alpha} > \mu_\alpha(y), \qquad \text{if } \tfrac{\pi}{4} < \alpha < \tfrac{\pi}{2}.$$

Therefore we have necessarily $y = 0$ or $\theta = 0$. Remark that $\phi_\alpha(y,0) = (\tau_\alpha(y))^2$ and $\phi_\alpha(0,\theta) = (\sigma_\alpha(\theta))^2$, therefore going back to (17) we get

$$C_{S_\alpha}(2) = \max\big(\sup\{\tau_\alpha(y)\,;\, y>0\}, \sup\{\sigma_\alpha(c)\,;\, c\in(0,1)\}\big). \tag{18}$$

But we have $\tau_\alpha(0) = \sigma_\alpha(0)$ and (see the next lemma) τ_α is decreasing for $\alpha > \frac{\pi}{4}$ and σ_α is decreasing for $\alpha < \frac{\pi}{4}$. We deduce the theorem from (18). □

LEMMA 13. *The functions ν_α, μ_α, τ_α and σ_α have the following properties*

a) μ_α is increasing for $\alpha < \pi/4$, decreasing for $\alpha > \pi/4$,
b) ν_α is decreasing for $\alpha < \pi/4$, increasing for $\alpha > \pi/4$,
c) τ_α is decreasing for $\alpha > \pi/4$,
d) σ_α is decreasing for $\alpha < \pi/4$.

Proof. a) The derivative μ'_α has the same sign as

$$\begin{aligned} f_\alpha(y) &:= \mathrm{ch}^2\tfrac{\pi}{2}y\ \mathrm{sh}^2 2\alpha y\ \mu'_\alpha(y) \\ &= \tfrac{\pi}{4}\,\mathrm{sh}\,4\alpha y\ - \alpha\,\mathrm{sh}\,\pi y\ - \cos 2\alpha\,\big(\tfrac{\pi}{2}\,\mathrm{sh}\,2\alpha y\ - \alpha\,\mathrm{sh}\,\pi y\ \mathrm{ch}\,2\alpha y\big) \end{aligned}$$

We clearly have $f_\alpha(0) = 0$ and

$$f'_\alpha(y) = \pi\,\alpha\cos 2\alpha\ g_\alpha(y), \quad \text{where}$$

$$g_\alpha(y) := \tfrac{\mathrm{ch}\,4\alpha y - \mathrm{ch}\,\pi y}{\cos 2\alpha} + \mathrm{ch}\,2\alpha y\,(\mathrm{ch}\,\pi y - 1) + \tfrac{2\alpha}{\pi}\,\mathrm{sh}\,\pi y\ \mathrm{sh}\,2\alpha y,$$

$$= -\tfrac{2\,\mathrm{sh}(\frac{\pi}{2}-2\alpha)y\ \ \mathrm{sh}(\frac{\pi}{2}+2\alpha)y}{\cos 2\alpha} + 2\,\mathrm{sh}\,\tfrac{\pi}{2}y\,\big(\mathrm{sh}\,\tfrac{\pi}{2}y\ \mathrm{ch}\,2\alpha y\ + \tfrac{2\alpha}{\pi}\,\mathrm{ch}\,\tfrac{\pi}{2}y\ \mathrm{sh}\,2\alpha y\big)$$

$$= \mathrm{sh}\,\tfrac{\pi y}{2}\,\Big[(1+\tfrac{2\alpha}{\pi})\ \mathrm{sh}\,(\tfrac{\pi}{2}+2\alpha)y + (1-\tfrac{2\alpha}{\pi})\ \mathrm{sh}\,(\tfrac{\pi}{2}-2\alpha)y - 2\tfrac{\mathrm{sh}(\frac{\pi}{2}-2\alpha)y\ \mathrm{sh}(\frac{\pi}{2}+2\alpha)y}{\mathrm{sh}\,\frac{\pi}{2}y\ \cos 2\alpha}\Big].$$

We shall prove that $\forall y > 0$, $\forall\alpha \in (0,\frac{\pi}{2})$, we have $g_\alpha(y) > 0$, thus μ'_a has the same sign as $\cos 2\alpha$ and the part a) of the lemma follows. (Note that $g_0(y) = g_{\pi/2}(y) = 0$ and $g_{\pi/4}(y) = \frac{3}{2}\,\mathrm{sh}\,\pi y\,(\mathrm{sh}\,\frac{\pi}{2}y - \frac{4}{3}y)$).

It is clear that $g_\alpha(y) > 0$ if $1 + \frac{2\alpha}{\pi} - 2\frac{\mathrm{sh}(\frac{\pi}{2}-2\alpha)y}{\mathrm{sh}\,\frac{\pi}{2}y\ \cos 2\alpha}\ \geq 0$; note that we have

$$\forall y > 0, \qquad \frac{|\,\mathrm{sh}(\frac{\pi}{2}-2\alpha)y|}{\mathrm{sh}\,\frac{\pi}{2}y} \leq \frac{|\pi/2-2\alpha|}{\pi/2},$$

therefore we have $g_\alpha(y) \geq 0$ as soon as $1 + \frac{2\alpha}{\pi} - 2\frac{|\pi/2-2\alpha|}{\pi/2\ \cos 2\alpha}\ \geq 0$, i.e. if we have the inequality $|\cos 2\alpha| \geq 2\frac{|\pi/2-2\alpha|}{\pi/2+\alpha}$. This is the case for $\frac{\pi}{6} \leq \alpha \leq \frac{\pi}{2}$. We turn now to the case $0 < \alpha < \frac{\pi}{6}$. Then we have

$$\frac{\operatorname{sh}(\frac{\pi}{2}-2\alpha)y\ \operatorname{sh}(\frac{\pi}{2}+2\alpha)y}{\operatorname{sh}^2\frac{\pi}{2}y} = \frac{\operatorname{sh}^2\frac{\pi}{2}y - \operatorname{sh}^2 2\alpha y}{\operatorname{sh}^2\frac{\pi}{2}y} < 1.$$

Therefore

$$g_\alpha(y) \geq \operatorname{sh}\frac{\pi y}{2}\Big[(1+\tfrac{2\alpha}{\pi})\operatorname{sh}(\tfrac{\pi}{2}+2\alpha)y + (1-\tfrac{2\alpha}{\pi})\operatorname{sh}(\tfrac{\pi}{2}-2\alpha)y - \frac{2}{\cos 2\alpha}\sqrt{\operatorname{sh}(\tfrac{\pi}{2}-2\alpha)y\ \operatorname{sh}(\tfrac{\pi}{2}+2\alpha)y}\ \Big].$$

We deduce $g_\alpha(y) > 0$ if $(\frac{1}{\cos 2\alpha})^2 \leq (1+\frac{2\alpha}{\pi})(1-\frac{2\alpha}{\pi})$, i.e. $\frac{2\alpha^2}{\pi^2} < \sin 4\alpha$ which is the case since $0 < \alpha < \frac{\pi}{6}$. A similar proof gives the part b).

c) The derivative τ_α' has the same sign as

$$\begin{aligned}
\operatorname{ch}^2\tfrac{\pi y}{4}\ \operatorname{sh}^2\alpha y\ \tau_\alpha'(y) &= \tfrac{\pi}{4}(\sin\alpha + \sqrt{\sin^2\alpha + \operatorname{sh}^2\alpha y})\operatorname{sh}\alpha y + \\
&\quad +\alpha\operatorname{sh}\tfrac{\pi y}{4}\operatorname{ch}\tfrac{\pi y}{4}\Big(\tfrac{\operatorname{sh}^2\alpha y\operatorname{ch}\alpha y}{\sqrt{\sin^2\alpha+\operatorname{sh}^2\alpha y}} - \sin\alpha\operatorname{ch}\alpha y - \operatorname{ch}\alpha y\sqrt{\sin^2\alpha+\operatorname{sh}^2\alpha y}\Big) \\
&= \tfrac{\sin\alpha+\sqrt{\sin^2\alpha+\operatorname{sh}^2\alpha y}}{\sqrt{\sin^2\alpha+\operatorname{sh}^2\alpha y}}\Big(\tfrac{\pi}{4}\operatorname{sh}\alpha y\sqrt{\sin^2\alpha+\operatorname{sh}^2\alpha y} - \tfrac{\alpha\sin\alpha}{2}\operatorname{sh}\tfrac{\pi y}{2}\operatorname{ch}\alpha y\Big).
\end{aligned}$$

Therefore τ_α' has the same sign as

$$\begin{aligned}
f_\alpha(y) &:= \tfrac{2\operatorname{sh}^2\alpha y\ (\sin^2\alpha+\operatorname{sh}^2\alpha y)}{\alpha^2\sin^2\alpha} - \tfrac{8}{\pi^2}\operatorname{sh}^2\tfrac{\pi y}{2}\operatorname{ch}^2\alpha y \\
&= \tfrac{\operatorname{ch}2\alpha y-1}{\alpha^2} + \tfrac{\operatorname{ch}4\alpha y-4\operatorname{ch}2\alpha y+3}{4\alpha^2\sin^2\alpha} - \tfrac{\operatorname{ch}(\pi+2\alpha)y+\operatorname{ch}(\pi-2\alpha)y+2\operatorname{ch}\pi y-2\operatorname{ch}2\alpha y-2}{\pi^2} \\
&= -\sum_{k\geq 2}\beta_k(\alpha)\frac{y^{2k}}{(2k)!},
\end{aligned}$$

with

$$\beta_k(\alpha) := \frac{(\pi+2\alpha)^{2k}+(\pi-2\alpha)^{2k}+2\,\pi^{2k}-2\,(2\alpha)^{2k}}{\pi^2} - \frac{(4\alpha)^{2k}-4(2\alpha)^{2k}\cos^2\alpha}{4\alpha^2\sin^2\alpha}.$$

Note that $\beta_1(\alpha) = 4-4 = 0$. The decreasing property of τ_α will follow from the inequalities

$$\beta_k(\alpha) > 0, \quad \text{for all } k \geq 2 \quad \text{and} \quad \beta \in (\frac{\pi}{4},\frac{\pi}{2}).$$

To prove them we first remark that $\beta_2(\alpha) = 4\big(\pi^2+8\alpha^2-\frac{12\alpha^2}{\sin^2\alpha}\big) > 0$, for $\alpha \in (\frac{\pi}{4},\frac{\pi}{2})$. For $k \geq 3$ we have

$$\beta_k(\alpha) \geq \frac{(\pi+2\alpha)^{2k}}{\pi^2} - \frac{(4\alpha)^{2k}}{4\alpha^2\sin^2\alpha} \geq (4\alpha)^{2k-6}\Big[\frac{(\pi+2\alpha)^6}{\pi^2} - \frac{(4\alpha)^6}{4\alpha^2\sin^2\alpha}\Big].$$

In order to complete the proof it is sufficient to show that

$$\gamma := (\pi+2\alpha)^6\sin^2\alpha - 1024\,\pi^2\,\alpha^4 > 0.$$

It is easily seen that for $\alpha \in (\frac{\pi}{4},\frac{\pi}{2})$ we have $\sin^2\alpha\,(\pi+2\alpha)^2 \geq 16\,\alpha^2$, therefore

$$\gamma \geq 16\,\alpha^2\left[(\pi+2\alpha)^4 - 64\,\pi^2\,\alpha^2\right] = 16\,\alpha^2\left[(\pi+2\alpha)^2+8\,\pi\,\alpha\right](\pi-2\alpha)^2 > 0.$$

A similar proof gives part d). □

From Theorem 9 and Theorem 12 it results by taking the limit as $\alpha \to 0$

COROLLARY 14. *In the two-dimensional case we have for the strip domain*

$$C_B(2) = \max_{y>0}\frac{1+\sqrt{1+y^2}}{y}\ \operatorname{th}\frac{\pi y}{4} \simeq 1.58765981...$$

References

[1] T. Ando, Structure of operators with numerical range 1, *Acta Sci. Math. (Szeged)* **34** (1973) 11–15.

[2] C. Badea, Personal communication.

[3] C. A. Berger, J. G. Stampell, Mapping theorems for the numerical range, *Amer. J. Math.* **89** (1967) 1047–1055.

[4] M. Crouzeix, B. Delyon, Some estimates for analytic functions of band or sectorial operators, *submitted* (2002).

[5] B. Delyon, F. Delyon, Generalization of Von Neumann's spectral sets and integral representation of operators, *Bull. Soc. Math. Fr.* **127** no. 1 (1999) 25–42.

[6] K. E. Gustafson, D. K. M. Rao, *Numerical Range*, Universitext, Springer-Verlag, (1997).

[7] P. Henrici, *Applied and computational complex analysis*, Wiley-Interscience, vol. 3, (1985).

[8] R. Nevanlinna, Über beschränkte Funktionen, die in gegebenen Punkten vorgeschriebene Werte annehmen, *Ann. Acad. Sci. Fenn. Ser. A, no. 1* **13** (1919).

[9] G. Pick, Über die Beschränkungen analytischer Funktionen, welche durch vorgebene Funktionwerte bewirkt werden, *Math. Ann.* **77** (1916) 7–23.

[10] D. Sarason, Generalized interpolation in H^∞, *Trans. Amer. Math. Soc.* **127** (1967) 179–203.

[11] J. von Neumann, Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes, *Math. Nachrichten* **4** (1951) 258–281.

[12] J. L .Walsh, Interpolation and Approximation by Rational Functions in the Complex Domain, *Amer. Math. Soc. Colloq. Publ. Providence, RI* **20** (1956).

Institut de Recherche Mathématique de Rennes, UMR 6625,
Université de Rennes 1, Campus de Beaulieu
35042-Rennes Cedex, France

E-mail address : michel.crouzeix@univ-rennes1.fr

Mathematics subject classification: 47 A 12