Numerical range and functional calculus in Hilbert space

Michel Crouzeix

Abstract

We prove an inequality related to polynomial functions of a square matrix, involving the numerical range of the matrix. We also show extensions valid for bounded and also unbounded operators in Hilbert spaces, which allow the development of a functional calculus.

Mathematical subject classification: 47A12 ; 47A25

Keywords: functional calculus, numerical range, spectral sets.

1 Introduction

In this paper H denotes a complex Hilbert space equipped with the inner product $\langle .\,, .\rangle$ and corresponding norm $\|v\| = \langle v, v\rangle^{1/2}$. We denote its unit sphere by $\Sigma_H := \{v \in H; \|v\| = 1\}$. For a bounded linear operator $A \in \mathcal{L}(H)$ we use the operator norm $\|A\|$ and denote by $W(A)$ its numerical range:

$$\|A\| := \sup_{v\in\Sigma_H} \|A\,v\|, \qquad W(A) := \{\langle A\,v, v\rangle\,;\, v \in \Sigma_H\}.$$

Recall that the numerical range is a convex subset of $\mathbb{C}$ and that its closure contains the spectrum of A. We keep the same notation in the particular case where $H = \mathbb{C}^d$ and A is a $d \times d$ matrix.

The main aim of this article is to prove that the inequality

$$\|p(A)\| \le 56 \sup_{z\in W(A)} |p(z)| \tag{1}$$

holds for any matrix $A \in \mathbb{C}^{d,d}$ and any polynomial $p : \mathbb{C} \to \mathbb{C}$ (independently of d).

It is remarkable that the completely bounded version of this inequality is also valid. More precisely, we consider now matrix-valued polynomials $P : \mathbb{C} \to \mathbb{C}^{m,n}$, i.e., for $z \in \mathbb{C}$, $P(z) = (p_{ij}(z))$ is a matrix, with each entry $p_{ij}(.)$ a polynomial $\mathbb{C} \to \mathbb{C}$. Then the inequality

$$\|P(A)\| \le 56 \sup_{z\in W(A)} \|P(z)\|, \tag{2}$$

holds for any matrix $A \in \mathbb{C}^{d,d}$, any polynomial $P : \mathbb{C} \to \mathbb{C}^{m,n}$, and any values of d, m, n.

We now fix a matrix $A \in \mathbb{C}^{d,d}$ and consider the algebra $\mathbb{C}[z]$ of polynomials p from $W(A)$ into $\mathbb{C}$, provided with the norm $\|p\|_{\infty,A} = \sup_{z\in W(A)} |p(z)|$. The map $p \mapsto p(A)$ is then a homomorphism from the algebra $\mathbb{C}[z]$ into the algebra $\mathbb{C}^{d,d}$. The inequality (1) means that this homomorphism is bounded with constant 56, and the inequality (2) that it is complely bounded with constant 56. These inequalities may be summarized in the following statement.

Theorem 1. *For any matrix $A \in \mathbb{C}^{d,d}$ the homomorphism $p \mapsto p(A)$ from the algebra $\mathbb{C}[z]$, with norm $\|\,.\,\|_{\infty,A}$, into the algebra $\mathbb{C}^{d,d}$ is completely bounded, with constant* 56.

It will be clear from the proof that the value 56 is not optimal. Denoting by $\mathcal{Q}$ and $\mathcal{Q}_{cb}$ the smallest constants such that the inequalities

$$\|p(A)\| \leq \mathcal{Q} \sup_{z \in W(A)} \|p(z)\|, \qquad \|P(A)\| \leq \mathcal{Q}_{cb} \sup_{z \in W(A)} \|P(z)\|, \tag{3}$$

hold for any matrix $A \in \mathbb{C}^{d,d}$, any polynomial $p : \mathbb{C} \to \mathbb{C}$, $P : \mathbb{C} \to \mathbb{C}^{m,n}$ and any values of d, m, n, we have

$$2 \leq \mathcal{Q} \leq \mathcal{Q}_{cb} \leq 56.$$

The lower bound is attained by the polynomial $p(z) = z$ and the matrix $A = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$, in which case $W(A)$ is the unit disk.

Using the fact that these constants are independent of d, we will deduce that the inequalities (3) are still valid for any bounded linear operator $A \in \mathcal{L}(H)$ on any Hilbert space H. For a convex subset E of $\mathbb{C}$ we introduce the algebra $\mathcal{H}_b(E)$ of continuous and bounded functions in $\overline{E}$ which are holomorphic in the interior of E. The following statement shows the existence of a functional calculus based on the numerical range.

Theorem 2. *Let H be a Hilbert space. For any bounded linear operator $A \in \mathcal{L}(H)$ the homomorphism $p \mapsto p(A)$ from the algebra $\mathbb{C}[z]$, with norm $\|\,.\,\|_{\infty,A}$, into the algebra $\mathcal{L}(H)$, is bounded with constant $\mathcal{Q}$. It admits a unique bounded extension from $\mathcal{H}_b(W(A))$ into $\mathcal{L}(H)$. This extension is bounded with constant $\mathcal{Q}$ and completely bounded with constant $\mathcal{Q}_{cb}$.*

For many applications it is useful to consider unbounded operators. Assume now that $A \in \mathcal{L}(D(A), H)$ is a closed linear operator with domain $D(A)$ densely and continuously embedded in H. Its numerical range is then defined by $W(A) := \{\langle A\,v, v\rangle \,; v \in \Sigma_H \cap D(A)\}$. We assume that the spectrum $\sigma(A)$ is included in $\overline{W(A)}$. The numerical range is still convex but unbounded and we hence cannot apply polynomials. We therefore instead consider the algebra $\mathbb{C}_b(z)$ of rational functions which are bounded on $W(A)$, and provide it with the norm $\|r\|_{\infty,A} = \sup_{z \in W(A)} \|r(z)\|$. We then have the following statement similar to Theorem 2

Theorem 3. *For any closed linear unbounded operator A such that $\sigma(A) \subset \overline{W(A)}$, the homomorphism $r \mapsto r(A)$ from the algebra $\mathbb{C}_b(z)$, with norm $\|\,.\,\|_{\infty,A}$, into the algebra $\mathcal{L}(H)$, is bounded with constant $\mathcal{Q}$. It admits a unique bounded extension from the algebra $\mathcal{H}_b(W(A))$ into $\mathcal{L}(H)$. This extension is bounded with constant $\mathcal{Q}$ and completely bounded with constant $\mathcal{Q}_{cb}$.*

An outline of this paper is the following : in Section 2 we deduce Theorems 2 and 3 from the inequality (3), and illustrate them by application to the theory of Cosine functions. Section 3 is devoted to an integral representation formula for $P(A)$, which is the key to proving (2). In Section 4 we prove some preliminary bounds which allow us to treat most matrices. The remaining case, which corresponds to very thin numerical ranges, is more complicated and is considered in Section 5. Section 6 is devoted to the proof of a geometrical inequality for convex domains in the plane, which is needed in Section 4. Some concluding remarks are given in Section 7.

The author conjectures that, in fact, $\mathcal{Q} = \mathcal{Q}_{cb} = 2$.

2 The extension to infinite dimension

Proof of Theorem 2. We assume that (3) is valid if $H = \mathbb{C}^d$, for all integers d. This implies that (3) holds for any finite dimensional Hilbert space. Let now H be infinite-dimensional and consider an operator $A \in \mathcal{L}(H)$ and a polynomial p of degree n. For $u \in H$ given, we introduce the Krylov space $\mathcal{K}_n = \operatorname{Span}\{u, Au, \dots, A^n u\} \subset H$, denote by Π_n the orthogonal projection onto $\mathcal{K}_n$, and set $A_n := \Pi_n A|_{\mathcal{K}_n}$. We clearly have $p(A)\,u = p(A_n)\,u$ and $W(A_n) \subset W(A)$, and A_n acts on the $n{+}1$ dimensional space $\mathcal{K}_n$. We then deduce that

$$\|p(A)\,u\| = \|p(A_n)\,u\| \le \mathcal{Q} \sup_{z \in W(A_n)} |p(z)|\ \|u\| \le \mathcal{Q} \sup_{z \in W(A)} |p(z)|\ \|u\|,$$

which implies $\|p(A)\| \le \mathcal{Q} \sup_{z \in W(A)} |p(z)|$. The same proof works also for polynomials with matrix values, thus the inequalities (3) are still valid.

The Mergelyan theorem states that $\mathbb{C}[z]$ is dense in $\mathcal{H}_b(W(A))$, which completes the proof. □

Proof of Theorem 3. We consider now a closed linear unbounded operator A such that $\sigma(A) \subset \overline{W(A)}$ and a rational function r bounded on $\overline{W(A)}$. It is clear that $r(A)$ is well defined and belongs to $\mathcal{L}(H)$ (write r in simple partial fraction form). In the case $W(A) = \mathbb{C}$, the space $\mathcal{H}_b(W(A))$ is reduced to constant functions and Theorem 2 follows immediately (but has no interest). In the other cases, changing, if needed, A to $\alpha + e^{i\theta}A$, $\alpha \in \mathbb{C}$, $\theta \in \mathbb{R}$, we have only two possibilities to consider :

a) The numerical range is a strip : $\overline{W(A)} = \{z \in \mathbb{C}\,; 0 \le \operatorname{Re} z \le a\}$, $a > 0$.
We then set $A_\varepsilon = (1{+}\varepsilon A^*)^{-1} A (1{+}\varepsilon A)^{-1}$, for $\varepsilon > 0$.

b) $\{x\,; x > 0\} \subset W(A) \subset \{z \in \mathbb{C}\,; \operatorname{Re} z \ge 0\}$.
We then set $A_\varepsilon = A(1{+}\varepsilon A)^{-1}$, for $\varepsilon > 0$.

In both cases it is easy to verify that A_ε is a bounded operator and that $W(A_\varepsilon) \subset W(A)$. Let r be a rational function bounded in $W(A)$, clearly $r \in \mathcal{H}_b(W(A_\varepsilon))$. From Theorem 2 we have

$$\|r(A_\varepsilon)\| \le \mathcal{Q} \sup_{z \in W(A_\varepsilon)} |r(z)| \le \mathcal{Q} \sup_{z \in W(A)} |r(z)|. \tag{4}$$

We now note that $\lim_{\varepsilon \to 0} r(A_\varepsilon) = r(A)$, strongly on H. Indeed, it suffices to verify that

$$\lim_{\varepsilon \to 0} (\lambda - A_\varepsilon)^{-1} u = (\lambda - A)^{-1} u \quad \text{ for all } \lambda \notin \overline{W(A)} \text{ and all } u \in D(A).$$

We consider the case of a strip, the other situation being simpler. We have by a simple calculation

$$\begin{aligned}(\lambda{-}A_\varepsilon)^{-1}u - (\lambda{-}A)^{-1}u = &-\varepsilon(\lambda{-}A_\varepsilon)^{-1}(1{+}\varepsilon A^*)^{-1}(A{+}A^*)(1{+}\varepsilon A)^{-1}(\lambda{-}A)^{-1}Au\\ &-\varepsilon(\lambda{-}A_\varepsilon)^{-1}(1{+}\varepsilon A^*)^{-1}(\varepsilon A^*)(1{+}\varepsilon A)^{-1}A(\lambda{-}A)^{-1}Au.\end{aligned}$$

The convergence to 0 of the second member follows from the bounds

$$\|A + A^*\| \le 2\,a, \quad \|(1 + \varepsilon A^*)^{-1}\| \le 1, \quad \|(1 + \varepsilon A^*)^{-1}(\varepsilon A^*)\| \le 1,$$
$$\|(\lambda{-}A_\varepsilon)^{-1}\| \le \frac{1}{d(\lambda, W(A))}, \quad \|(\lambda{-}A)^{-1}\| \le \frac{1}{d(\lambda, W(A))},$$
$$\|A\,(\lambda{-}A)^{-1}\| \le \frac{|\lambda|}{d(\lambda, W(A))}.$$

Taking the limit as $\varepsilon \to 0$ in (4), we obtain that the map $r \mapsto r(A)$ is bounded with constant $\mathcal{Q}$. The space $\mathbb{C}_b(z)$ is dense in the subset $\mathcal{H}_{b,0}(W(A))$ of functions of $\mathcal{H}_b(W(A))$ which have a limit at ∞. This therefore allows a first extension of the map to this subspace.

Let us consider now $f \in \mathcal{H}_b(W(A))$ and $\lambda \notin \overline{W(A)}$. With $g_\lambda(z) = (\lambda - z)^{-1} f(z)$ we have $g_\lambda \in \mathcal{H}_{b,0}(W(A))$, and from the previous extension, $g_\lambda(A) \in \mathcal{L}(H)$ is well defined. If, in addition, $f \in \mathcal{H}_{b,0}(W(A))$, then $f(A) \in \mathcal{L}(H)$ and $g_\lambda(A) = (\lambda - A)^{-1} f(A)$. This shows that $g_\lambda(A) \in \mathcal{L}(H, D(A))$ and $f(A) = (\lambda - A) g_\lambda(A)$.

We return to the general case $f \in \mathcal{H}_b(W(A))$. Recall that we have assumed $W(A) \subset \{z \in \mathbb{C}\,; \operatorname{Re} z \geq 0\}$. From the inequality, with $\varepsilon > 0$,

$$\|A(1+\varepsilon A)^{-1} g_\lambda(A)\| \leq \mathcal{Q} \sup_{z \in W(A)} \left| \frac{z}{1+\varepsilon z} \frac{f(z)}{\lambda - z} \right| \leq \mathcal{Q}\, \frac{|\lambda|}{d(\lambda, W(A))}\, \|f\|_{\infty,A},$$

we deduce, by taking the limit as $\varepsilon \to 0$, that $\|A\, g_\lambda(A)\| \leq \mathcal{Q}\, \frac{|\lambda|}{d(\lambda, W(A))} \|f\|_{\infty,A}$ and $g_\lambda(A) \in \mathcal{L}(H, D(A))$. We can therefore define $f(A) \in \mathcal{L}(H)$ by $f(A) = (\lambda - A) g_\lambda(A)$. Passing to the limit in the inequality

$$\|(1+\varepsilon A)^{-1} f(A)\| \leq \mathcal{Q} \sup_{z \in W(A)} \left| \frac{\lambda - z}{1+\varepsilon z} \frac{f(z)}{\lambda - z} \right| \leq \mathcal{Q}\, \|f\|_{\infty,A},$$

we deduce $\|f(A)\| \leq \mathcal{Q}\, \|f\|_{\infty,A}$. It is easily seen that our definition does not depend on the particular choice of λ.

The same arguments work for the complete bound.

$\square$

In order to illustrate the previous results, we describe the following application, which has been mentioned previously in [3]. Assume that the unbounded operator A satisfies the hypotheses of Theorem 3, and that there exists $\alpha > 0$ such that

$$W(A) \subset \mathcal{P} := \{x + i\,y \in \mathbb{C}\,; y^2 \leq \alpha^2 x\}.$$

As a result of Theorem 3 the operators $\exp(it\sqrt{A})$ and $\cos(t\sqrt{A})$ are well defined for all $t \in \mathbb{R}$ and strongly continuous in H. The problem

$$u''(t) + A\,u(t) = 0, \quad \text{for } t \in \mathbb{R}, \qquad u(0) = u_0, \quad u'(0) = v_0,$$

with $u_0 \in D(A)$, $v_0 \in D(A^{1/2})$ given, therefore admits a unique solution $u \in C^2(\mathbb{R}; H) \cap C^0(\mathbb{R}; D(A))$, which may be written

$$u(t) = \cos(t\sqrt{A})\, u_0 + \int_0^t \cos(s\sqrt{A})\, v_0\, ds.$$

Examples of such operators A can be found among the ones associated with convection diffusion equations

$$Av := -\mathrm{div}\,(a\,\vec{\nabla}\,v) + \vec{b}.\vec{\nabla}v + cv,$$

with suitable assumptions on the coefficients a, $\vec{b}$ and c, and appropriate choices of the spaces H and $D(A)$. To our knowledge, until now no other methods have provided such an existence result.

3 The representation formula

From now on, A will be a $d\times d$ matrix . Let Ω be a bounded convex domain of $\mathbb{C}$ such that $W(A)\subset\Omega$; we assume that its boundary $\partial\Omega$ is smooth and has a strictly positive curvature at each point. We then deduce from (2) that

$$\|P(A)\|\leq \mathcal{Q}_{cb}\sup_{z\in\Omega}\|P(z)\|,\qquad \text{for any matrix-valued polynomial } P. \tag{5}$$

Conversely, if this estimate holds for all such convex domains, we easily obtain (2).

After the use of a similarity transformation, we can assume that there exist two functions $\eta_\pm\in C^0([0,2])\cap C^\infty((0,2))$ such that

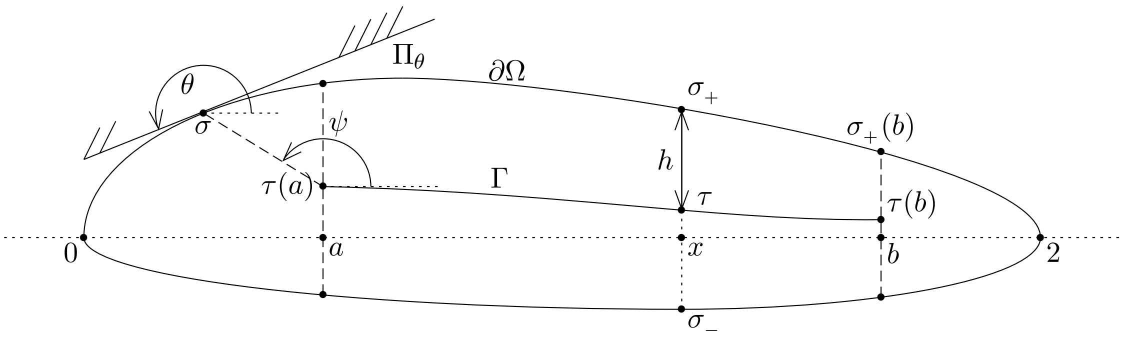
$$0,2\in\partial\Omega,\qquad \operatorname{diam}(\Omega)=2,\quad \Omega=\{x+i\,y\,;0<x<2,\ \eta_-(x)<y<\eta_+(x)\}.$$

We set $h(x):=\frac12(\eta_+(x)-\eta_-(x))$.

We will distinguish two possible cases, which we refer to as 'thin' and 'thick'.

(i) Thin domains : $h(1)<10^{-5}$.
Then there exist unique a, b such that $0<a<1<b<2$ and $h(a)/a=h(b)/(2-b)=1/6$. We introduce the median curve

$$\Gamma=\{z=\tau(x)\,;a\leq x\leq b,\ \tau(x)=x+\frac{i}{2}\,(\eta_+(x)+\eta_-(x))\}.$$



In order to make the figure more clear, we have enlarged the scale in the vertical direction.

(ii) Thick domains : $h(1)\geq 10^{-5}$. In this case we will use $a=b=1$.

We denote by σ the generic point of arclength s on the boundary $\partial\Omega$, oriented counterclockwise, and set $\theta=\arg(\frac{d\sigma}{ds})$ so that $\frac{d\sigma}{ds}=e^{i\theta}$. Due to the strict convexity, the point σ can be considered as a C^∞ function of θ. We define the following functions of the real variable x :

$$\begin{aligned}
&\sigma_+:(0,c)\to\mathbb{C}\ \text{ is defined by }\ \sigma_+(x)\in\partial\Omega,\operatorname{Im}\sigma_+(x)>0\text{ and }x=\operatorname{Re}\sigma_+(x),\\
&\sigma_-:(0,c)\to\mathbb{C}\ \text{ is defined by }\ \sigma_-(x)\in\partial\Omega,\operatorname{Im}\sigma_-(x)<0\text{ and }x=\operatorname{Re}\sigma_-(x),\\
&\theta_+=\arg(\frac{d\sigma_+}{ds}),\qquad \theta_-=\arg(\frac{d\sigma_-}{ds}).
\end{aligned}$$

We also use the following notations

$$\mu(\sigma,z)=\frac1\pi\,\frac{d}{ds}\big(\arg(\sigma-z)\big)=\frac{1}{2\pi i}\,(\frac{e^{i\theta}}{\sigma-z}-\frac{e^{-i\theta}}{\bar\sigma-\bar z}),$$

$$\Pi_\theta=\{z\,;\mu(\sigma,z)>0\}=\{z\,;\operatorname{Re}\frac{e^{i\theta}}{i}\,(\bar\sigma-\bar z)>0\}.$$

$$
\begin{aligned}
\xi(\theta,t) &= t\sigma(\theta)+(1-t)\,\tau(a) && \text{if}\quad \operatorname{Re}\sigma(\theta)\le a,\\
&= t\sigma(\theta)+(1-t)\,\tau(\operatorname{Re}\sigma(\theta)) && \text{if}\quad \operatorname{Re}\sigma(\theta)\in[a,b],\\
&= t\sigma(\theta)+(1-t)\,\tau(b) && \text{if}\quad \operatorname{Re}\sigma(\theta)\ge b,
\end{aligned}
$$

$$
\begin{aligned}
Jp(\theta,\bar z) &= \frac{1}{\pi}\int_0^1 p(\xi(\theta,t))\,\frac{\partial\xi}{\partial t}\,\frac{e^{2i\theta}(\bar\sigma(\theta)-\bar z)}{(\xi(\theta,t)-\sigma(\theta)+e^{2i\theta}(\bar\sigma(\theta)-\bar z))^2}\,dt,\\
G_+p(\bar z) &= \frac{1}{2\pi i}\int_a^b \frac{p(\tau)\,e^{-2i\theta_+}}{\bar\tau-\bar z-i\,(1+e^{-2i\theta_+})\,h}\,\frac{d\tau}{dx}\,dx,\\
G_-p(\bar z) &= \frac{1}{2\pi i}\int_a^b \frac{p(\tau)\,e^{-2i\theta_-}}{\bar\tau-\bar z+i\,(1+e^{-2i\theta_-})\,h}\,\frac{d\tau}{dx}\,dx.
\end{aligned}
$$

The proof of our main result, Theorem 1, is based on the following lemma.

Lemma 4. *For any polynomial p and any $z\in\overline{\Omega}$, we have the representation formula*

$$
p(z)=\int_{\partial\Omega}p(\sigma)\,\mu(\sigma,z)\,ds+\int_0^{2\pi}Jp(\theta,\bar z)\,d\theta-G_+p(\bar z)+G_-p(\bar z). \tag{6}
$$

Furthermore,

$$
\sup_{z\in\Pi_\theta}|Jp(\theta,\bar z)|\le(\tfrac12+|\cot(\theta-\psi)|)\sup_{\zeta\in\Omega}|p(\zeta)|, \tag{7}
$$

where

$$
\psi=\begin{cases}\arg(\sigma-\tau(a)) & \textit{if } \operatorname{Re}\sigma\le a,\\ \arg(\sigma-\bar\sigma)) & \textit{if } a<\operatorname{Re}\sigma<b,\\ \arg(\sigma-\tau(b)) & \textit{if } \operatorname{Re}\sigma\ge b,\end{cases}
$$

defined in such a way that $\theta-\psi\in(0,\pi)$.

Proof. From the Cauchy formula we have

$$
p(z)=\frac{1}{2\pi i}\int_{\partial\Omega}p(\sigma)\,\frac{d\sigma}{\sigma-z}.
$$

The relation (6) is therefore equivalent to

$$
\frac{1}{2\pi i}\int_{\partial\Omega}p(\sigma)\,\frac{d\bar\sigma}{\bar\sigma-\bar z}=\int_0^{2\pi}Jp(\theta,\bar z)\,d\theta-G_+p(\bar z)+G_-p(\bar z). \tag{8}
$$

We first note that $Jp(\theta,\bar z)$ is well defined, continuous and analytic in $\bar z$, for all $\theta\in\mathbb{R}$ and all $z\in\overline{\Omega}$. We set

$$
v_1(\theta,t)=-\frac{1}{2\pi i}\,\frac{p(\xi(\theta,t))}{\xi(\theta,t)-\sigma+e^{2i\theta}(\bar\sigma-\bar z)}\,\frac{\partial\xi}{\partial t},\quad v_2(\theta,t)=\frac{1}{2\pi i}\,\frac{p(\xi(\theta,t))}{\xi(\theta,t)-\sigma+e^{2i\theta}(\bar\sigma-\bar z)}\,\frac{\partial\xi}{\partial\theta}.
$$

Using that $\frac{\partial}{\partial\theta}\big(-\sigma+e^{2i\theta}(\bar\sigma-\bar z)\big)=2\,i\,e^{2i\theta}(\bar\sigma-\bar z)$, we get

$$
\operatorname{div}\vec v=\frac{\partial v_1}{\partial\theta}+\frac{\partial v_2}{\partial t}=\frac{1}{\pi}\,\frac{\partial\xi}{\partial t}\,p(\xi(\theta,t))\,\frac{e^{2i\theta}(\bar\sigma-\bar z)}{(\xi(\theta,t)-\sigma+e^{2i\theta}(\bar\sigma-\bar z))^2},
$$

and hence by Green's formula,

$$
\int_0^{2\pi}Jp(\theta,z)\,d\theta \;=\; \int_0^{2\pi}\!\!\int_0^1\operatorname{div}\vec v\;dt\,d\theta=\int_0^{2\pi}v_2(\theta,1)\,d\theta-\int_{\pi/2}^{5\pi/2}v_2(\theta,0)\,d\theta.
$$

We clearly have

$$\int_0^{2\pi} v_2(\theta,1)\,d\theta = \frac{1}{2\pi i}\int_{\partial\Omega}\frac{p(\sigma)}{e^{2i\theta}(\bar\sigma-\bar z)}\,d\sigma = \frac{1}{2\pi i}\int_{\partial\Omega}p(\sigma)\,\frac{d\bar\sigma}{\bar\sigma-\bar z},$$

$$\int_{\pi/2}^{3\pi/2} v_2(\theta,0)\,d\theta \;=\; \frac{1}{2\pi i}\int_{\theta_+(b)}^{\theta_+(a)}\frac{p(\tau)}{\tau-\sigma+e^{2i\theta}(\bar\sigma-\bar z)}\,\frac{d\tau}{d\theta}\,d\theta,$$

and thus, since $\overline{\sigma}_+-\bar\tau=\tau-\sigma_+=-i\,h$,

$$\int_{\pi/2}^{3\pi/2} v_2(\theta,0)\,d\theta = \frac{1}{2\pi i}\int_b^a \frac{p(\tau)\,e^{-2i\theta_+}}{\bar\tau-\bar z-i\,(1+e^{-2i\theta_+})\,h}\,d\tau = -G_+p(\bar z).$$

Similarly,

$$\int_{3\pi/2}^{5\pi/2} v_2(\theta,0)\,d\theta = G_-p(\bar z),$$

which shows (8).

It remains to prove (7). From the maximum principle we have, if, for instance, $a\le \mathrm{Re}(\sigma)\le b$, that, for $z\in\Pi_\theta$,

$$\begin{aligned}
|Jp(\theta,\bar z)| &\le \sup_{y\in\mathbb{R}}|Jp(\theta,\sigma+ye^{i\theta})|\\
&\le \sup_{y\in\mathbb{R}}\int_0^1 |p(\xi(\theta,t))|\,\frac{|\frac{\sigma-\bar\sigma}{2}||y|}{\pi\,|(t-1)\frac{\sigma-\bar\sigma}{2}+e^{i\theta}y|^2}\,dt\\
&\le \sup_{y\in\mathbb{R}}\int_0^\infty \frac{|\frac{\sigma-\bar\sigma}{2}||y|}{\pi\,|s\frac{\sigma-\bar\sigma}{2}+e^{i\theta}y|^2}\,ds\quad \sup_{\zeta\in\Omega}|p(\zeta)|\\
&\le \max_{\varepsilon=\pm1}\int_0^\infty \frac{d\tau}{\pi\,|\tau e^{i\psi}+\varepsilon\,e^{i\theta}|^2}\ \sup_{\zeta\in\Omega}|p(\zeta)| = \frac{\max(\theta-\psi,\pi-\theta+\psi)}{\pi\,\sin(\theta-\psi)}\ \sup_{\zeta\in\Omega}|p(\zeta)|.
\end{aligned}$$

The estimate (7) then follows from the inequality $\max\,(u,\pi-u)\le\frac{\pi}{2}\sin u+\pi|\cos u|$, which holds for $u=\theta-\psi\in(0,\pi)$. The other possible cases for σ may be treated in the same way. ☐

We consider now a square matrix A satisfying $W(A)\subset\Omega$. Then the operator

$$\mu(\sigma,A)=\frac{1}{2\pi i}\big(e^{i\theta}(\sigma-A)^{-1}-e^{-i\theta}(\bar\sigma-A^*)^{-1}\big)$$

is well defined and, it is easily verified that, for $v\in\mathbb{C}^d$, $\langle\mu(\sigma,A)\,v,v\rangle\ge0$. The operator $\mu(\sigma,A)$ is thus positive semi-definite.

Theorem 5. *Assume that $W(A)\subset\Omega$. Let P be a polynomial with values in $\mathbb{C}^{m,n}$. We then have*

$$P(A)=\int_{\partial\Omega}P(\sigma)\otimes\mu(\sigma,A)\,ds+\int_0^{2\pi}JP(\theta,A^*)\,d\theta-G_+P(A^*)+G_-P(A^*). \tag{9}$$

Proof. The relation (9) is a consequence of

$$P(A)=\frac{1}{2\pi i}\int_{\partial\Omega}P(\sigma)\otimes(\sigma-A)^{-1}\,d\sigma,$$

and

$$\frac{1}{2\pi i}\int_{\partial\Omega}P(\sigma)\otimes(\bar\sigma-A^*)^{-1}\,d\bar\sigma \;=\; \int_0^{2\pi}JP(\theta,A^*)\,d\theta-G_+P(A^*)+G_-P(A^*).$$

The first equality is the Cauchy formula and the second follows from the relation (8), i.e.,

$$\frac{1}{2i\pi}\int_{\partial\Omega} P(\sigma)\frac{d\bar\sigma}{\bar\sigma-\bar z} = \int_0^{2\pi} JP(\sigma,\bar z)\,d\theta - G_+P(\bar z) + G_-P(\bar z),$$

since each term in this relation is an analytic function of $\bar z$, for $z\in\Omega$. □

Remark. Note that the representation formula (9) is valid for thin as well as for thick domains Ω, but in the latter case we have the simplification $G_+P(A^*) = G_-P(A^*) = 0$.

4 Some preliminary bounds

For the proof of Theorem 1, we have to bound each term on the right of (9). The following lemma gives preliminary estimates for the first two terms.

Lemma 6. *Assume that the polynomial P satisfies $\|P(z)\|\le 1$, $\forall\, z\in\Omega$. Then we have*

$$\Big\|\frac{1}{2i\pi}\int_{\partial\Omega} P(\sigma)\otimes\mu(\sigma,A)\,ds\Big\| \le 2, \tag{10}$$

and

$$\Big\|\int_0^{2\pi} JP(\theta,A^*)\,d\theta\Big\| \le \pi + \int_0^{2\pi} |\cot(\theta-\psi)|\,d\psi. \tag{11}$$

(The relation between θ and ψ was defined in Lemma 4.)

Proof. a) Since the operator $\mu(\sigma,A)$ is positive, we have the inequality

$$\begin{aligned}\Big\|\int_{\partial\Omega} P(\sigma)\otimes\mu(\sigma,A)\,ds\Big\| &\le \Big\|\int_{\partial\Omega} \|P(\sigma)\|\ \mu(\sigma,A)\,ds\Big\|\\ &\le \Big\|\int_{\partial\Omega} \mu(\sigma,A)\,ds\Big\|\ \sup_{\zeta\in\Omega}\|P(\zeta)\|.\end{aligned}$$

We conclude the proof of (10) by noticing that, by the Cauchy formula,

$$\int_{\partial\Omega}\mu(\sigma,A)\,ds = 2I.$$

b) From the von Neumann inequality [5] for the half-plane Π_θ and the estimate (7), we have

$$\|JP(\theta,A^*)\| \le \sup_{z\in\Pi_\theta}|JP(\theta,\bar z)| \le \big(\tfrac12+|\cot(\theta-\psi)|\big)\ \sup_{\zeta\in\Omega}\|P(\zeta)\|.$$

Since

$$\int_{\partial\Omega}|\cot(\theta-\psi)|\,d(\theta-\psi) = 0,$$

we infer that

$$\|\int_0^{2\pi} JP(\theta,A^*)\,d\theta\| \le \int_0^{2\pi}(\frac12+|\cot(\theta-\psi)|)\,d\theta \le \pi + \int_0^{2\pi}|\cot(\theta-\psi)|\,d\psi.$$

□

The next lemma will complete the bound for the right hand side of (11) in the ‘thick’ case.

Lemma 7. *Assume that $h(1) \geq 10^{-5}$. Then we have*

$$\int_0^{2\pi} |\cot(\theta-\psi)|\, d\psi \leq 48.83.$$

Proof. a) Recall that in the 'thick' case $a = b = 1$. With $\rho(\psi) = |\sigma - \tau(a)|$, we have the polar representation $\sigma = \tau(a) + \rho e^{i\psi}$, and note that $\cot(\theta - \psi) = -\rho'/\rho$. Therefore,

$$\int_0^{2\pi} |\cot(\theta-\psi)|\, d\psi = \int_{\partial\Omega} \big|d \log |\sigma - \tau(a)|\big| = \mathrm{TV}(\log|\sigma-\tau(a)|).$$

We introduce the notations

$$R = \sup\{|\sigma-\tau(a)|; \sigma \in \partial\Omega\} \quad \text{and} \quad r = \inf\{|\sigma-\tau(a)|; \sigma \in \partial\Omega\}.$$

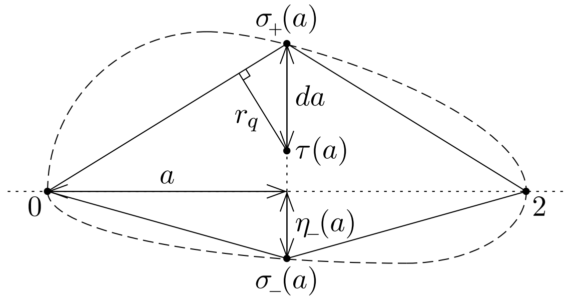
We conclude from Lemma 14, which will be stated and proved in Section 6, the estimate

$$\int_{\partial\Omega} \big|d \log |\sigma - \tau(a)|\big| \leq 2 \log \frac{R^4}{r^2(R^2-2r^2)}, \quad \text{if } \frac{R}{r} \geq 2, \quad (\leq 6\log 2 \quad \text{otherwise}).$$

b) We set $d = h(a)/a = h(1)$. Since the diameter is 2 we have $10^{-5} \leq d \leq 1$, and $R \leq 2\,a$. We also deduce from the convexity that Ω contains the quadrilateral with vertices

$0,\ \sigma_+(a),\ 2\,a,\ \sigma_-(a).$

Therefore the r corresponding to Ω is greater that the r_q which would correspond to this quadrilateral. The worst case corresponds to $\eta_-(a) = 0$, which gives $r \geq a\,d/\sqrt{4\,d^2+1}$, and therefore $R/r \leq 2\sqrt{4\,d^2+1}/d$.



Hence, by a simple calculation,

$$\int_{\partial\Omega} \big|d \log |\sigma - \tau(a)|\big| \leq 2 \max_{10^{-5}\leq d \leq 1} \log \frac{8(4d^2+1)^2}{d^2(7d^2+2)} \leq 48.83.$$

□

Recalling that the last two terms on the right of (9) vanish for thick domains. The two previous lemmas immediatly imply the following.

Corollary 8. *In the 'thick' case, i.e. when $h(1) \geq 10^{-5}$, we have*

$$\|P(A)\| \leq 54, \quad \textit{for all polynomials } P \textit{ with } \|P(z)\| \leq 1 \textit{ in } \Omega. \tag{12}$$

5 Bounds for a thin domain

We turn now to the 'thin' case. From the convexity of Ω we have, for $\sigma \in \partial\Omega$,

$$|\tan\theta| \leq \max\left\{\frac{|\operatorname{Im}\sigma|}{\operatorname{Re}\sigma}, \frac{|\operatorname{Im}\sigma|}{2-\operatorname{Re}\sigma}\right\} \qquad \text{and} \qquad |\operatorname{Im}\sigma| \leq 2\,h(\operatorname{Re}\sigma).$$

Recall that the numbers a, b with $a < b$ have been defined by $h(a)/a = 1/6 = h(b)/(2-b)$. This implies the following remark which will be frequently used in this section.

Remark. We have the estimate

$$|\tan\theta| \leq \frac{1}{3}, \qquad \textit{as soon as} \qquad a \leq \operatorname{Re}\sigma \leq b. \tag{13}$$

The next lemma will bound the right hand side of (11) in the 'thin' case.

Lemma 9. *We assume that $h(1) < 10^{-5}$. We then have*

$$\int_0^{2\pi} |\cot(\theta-\psi)|\,d\psi \le 7.551.$$

Proof. Similarly as in the previous lemma, we have, with $d = h(a)/a = 1/6$,

$$r_a := \inf\{|\sigma-\tau(a)|; \pi/2 < \psi < 3\pi/2\} \ge \frac{ad}{\sqrt{4d^2+1}} = \frac{a}{2\sqrt{10}}.$$

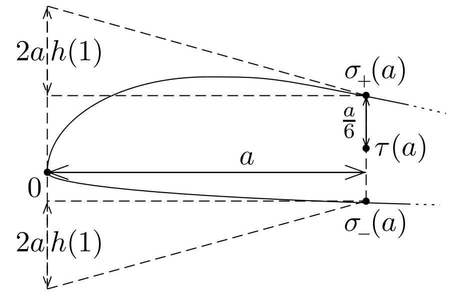
Using that $\tan\theta_+(a) \ge \tan\theta_+(1) \ge -2\,h(1)$, we deduce that the points $\sigma \in \partial\Omega$ such that $\psi \in (\pi/2, 3\pi/2)$ satisfy

$$0 \le \mathrm{Re}(\tau(a)-\sigma) \le a$$

and $\quad |\mathrm{Im}(\tau(a)-\sigma)| \le \dfrac{a}{6} + 2\,a\,h(1),$

which shows

$$R_a := \sup\{|\sigma-\tau(a)|; \frac{\pi}{2} < \psi < \frac{3\pi}{2}\} \le a\frac{\sqrt{38}}{6}.$$



In the same way we get for $\tau(b)$,

$$r_b := \inf_{\psi\in(-\frac{\pi}{2},\frac{\pi}{2})} |\sigma-\tau(b)| \ge \frac{b}{2\sqrt{10}} \quad \text{and} \quad R_b := \sup_{\psi\in(-\frac{\pi}{2},\frac{\pi}{2})} |\sigma-\tau(b)| \le b\frac{\sqrt{38}}{6}.$$

We now introduce the sets

$$\Omega_1 = \{z\in\Omega\,; \mathrm{Re}\,z \le a\} \cup \{\tau(a) + \frac{a}{b}(z-\tau(b))\,; z\in\Omega \text{ and } \mathrm{Re}\,z \ge b\},$$
$$\partial\Omega_a = \{\sigma\in\partial\Omega\,; \mathrm{Re}\,\sigma < a\}, \qquad \text{and} \qquad \partial\Omega_b = \{\sigma\in\partial\Omega\,; \mathrm{Re}\,\sigma > b\}.$$

It is easily seen that Ω_1 is convex and that

$$\begin{aligned}\int_0^{2\pi} |\cot(\theta-\psi)|\,d\psi &= \int_{\partial\Omega_a} \big|d\log|\sigma-\tau(a)|\big| + \int_{\partial\Omega_b} \big|d\log|\sigma-\tau(b)|\big|,\\ &= \int_{\partial\Omega_1} \big|d\log|\sigma-\tau(a)|\big|.\end{aligned}$$

Furthermore, we have

$$r_1 := \inf_{\sigma\in\partial\Omega_1} |\sigma-\tau(a)| \ge \frac{a}{2\sqrt{10}} \quad \text{and} \quad R_1 := \sup_{\sigma\in\partial\Omega_1} |\sigma-\tau(a)| \le a\frac{\sqrt{38}}{6}.$$

Applying Lemma 14 this yields

$$\int_0^{2\pi} |\cot(\theta-\psi)|\,d\psi \le 2\log\frac{380^2}{368\times 9} < 7.551.$$

□

We now write $A = B + i\,C$, with $B = \frac{1}{2}(A+A^*)$ and $C = \frac{1}{2i}(A-A^*)$ self-adjoint. Note that, from the assumption $h(1) \le 10^{-5}$, it follows that $\|C\| \le 2\,.\,10^{-5}$ and $\|B\| \simeq 2$, thus A is very close to B. We easily obtain a bound for $\|P(B)\|$ by spectral theory. The following lemma uses this for estimating $\|P(A)\|$.

Lemma 10. *Assume that the polynomial P satisfies $\|P(z)\| \leq 1$, $\forall\, z \in \Omega$. Then we have the estimate*

$$\|P(A)\| \leq 26.4 + \|G_+ P(A^*) - G_+ P(B)\| + \|G_- P(A^*) - G_- P(B)\|.$$

Proof. We deduce from (9), applied to both A and B,

$$\begin{aligned} P(A) - P(B) = \int_{\partial\Omega} p(\sigma) \otimes \mu(\sigma, A)\, ds - \int_{\partial\Omega} P(\sigma) \otimes \mu(\sigma, B)\, ds \\ + \int_0^{2\pi} JP(\theta, A^*)\, d\theta - \int_0^{2\pi} JP(\theta, B)\, d\theta \\ -G_+ P(A^*) + G_+ P(B) + G_- P(A^*) - G_- P(B). \end{aligned}$$

From spectral theory we have $\|P(B)\| \leq 1$, and, from Lemma 6 and Lemma 9,

$$\Big\| \int_{\partial\Omega} P(\sigma) \otimes \mu(\sigma, A)\, ds \Big\| \leq 2, \qquad \Big\| \int_{\partial\Omega} P(\sigma) \otimes \mu(\sigma, B)\, ds \Big\| \leq 2,$$
$$\Big\| \int_0^{2\pi} JP(\theta, A^*)\, d\theta \Big\| \leq 10.7, \qquad \Big\| \int_0^{2\pi} JP(\theta, B)\, d\theta \Big\| \leq 10.7.$$

The result stated now follows from the triangle inequality. $\square$

The following is the key tool for explaining how the numerical range enters into the proof.

Lemma 11. *Let*

$$\Gamma(\lambda, x) := h'(x)\,(\lambda - x) + h(x) \quad \textit{and} \quad \gamma(x) := \operatorname{Im} \tau(x), \qquad \textit{for} \ \ x \in (0, 2).$$

Then the condition $W(A) \subset \overline{\Omega}$ is equivalent to : $\quad 0 \leq B \leq 2$ *and*

$$|\langle (C - \gamma'(x)\,(B - x) - \gamma(x)) u, v \rangle| \leq \|(\Gamma(B, x))^{1/2} u\| \; \|\Gamma(B, x))^{1/2} v\|, \quad \forall\, u, v \in H, \forall\, x \in (0, 2).$$

This condition implies

$$|\langle C\, u, v \rangle| \leq 2\, \|(\Gamma(B, x))^{1/2} u\| \; \|\Gamma(B, x))^{1/2} v\|, \quad \forall\, u, v \in H.$$

Proof. a) We first remark that $X + iY \in \overline{\Omega}$ is equivalent to

$$0 \leq X \leq 2 \quad \text{and} \quad |Y - \gamma'(x)\,(X - x) - \gamma(x)| \leq \Gamma(X, x), \quad \forall\, x \in (0, 2).$$

Indeed, $Y - \gamma'(X - x) - \gamma(x) = \pm\Gamma(X, x)$ are the equations of the tangents to $\partial\Omega$ at the points σ_+ and σ_-. Therefore the condition $W(A) \subset \overline{\Omega}$ means $0 \leq B \leq 2$ and

$$|\langle (C - \gamma'(x)(B - x) - \gamma(x)) v, v \rangle| \leq \langle \Gamma(B, x)\, v, v \rangle, \qquad \forall v \in H, \ \forall\, x \in (0, 2).$$

Thus $\Gamma(B, x) \geq 0$ and the lemma follows from the generalized Cauchy-Schwarz inequality.

b) We note that, if $X \in (0, 2)$, then

$$\begin{aligned} \Gamma(X, x) - \gamma'(x)\,(X - x) - \gamma(x) &= -\eta_-'(x)\,(X - x) - \eta_-(x) > 0, \\ \Gamma(X, x) + \gamma'(x)\,(X - x) + \gamma(x) &= +\eta_+'(x)\,(X - x) + \eta_+(x) > 0. \end{aligned}$$

Therefore $X + iY \in \overline{\Omega}$ implies $|Y| \leq 2\,\Gamma(X, x)$. The proof is completed as in part a). $\square$

We now introduce the operator, for $t \in [0, 1]$, $x \in [a, b]$,

$$\begin{aligned} \Psi(t, x) \ &:= \ (\Gamma(B, x))^{1/2} (\bar{\tau} - i\,(1 + e^{-2i\theta_+})\, h - B + i\, t\, C)^{-1} \\ &= \ (\Gamma(B, x))^{1/2} (x - i\eta_+ - i\, h\, e^{-2i\theta_+} - B + i\, t\, C)^{-1}. \end{aligned} \tag{14}$$

The following technical lemma will be used to bound the third term on the right in (9).

Lemma 12. *We have, for all $u \in H$,*

$$\|\Psi(0,.)\,u\|_{L^2(a,b)} \le 1.33\sqrt{\pi}\,\|u\|, \qquad \|\Psi(1,.)\,u\|_{L^2(a,b)} \le 10.93\sqrt{\pi}\,\|u\|.$$

Proof. a) For brevity we will omit the subscript $+$ in the proof. Since C does not appear in Ψ for $t=0$, we deduce from spectral theory

$$\|\Psi(0,.)\,u\|_{L^2(a,b)} \le K_1\,\|u\|, \quad \text{with} \quad K_1 = \sup_{\lambda\in(0,2)} \Big(\int_a^b \frac{\Gamma(\lambda,x)}{|x-\lambda-i\,\eta-i\,h\,e^{-2i\theta}|^2}\,dx\Big)^{1/2}.$$

We have

$$\begin{aligned}
|x-\lambda-i\,\eta-i\,h\,e^{-2i\theta}|^2 &= (x-\lambda-h\,\sin 2\theta)^2 + (\eta + h\,\cos 2\theta)^2 \\
&\ge (x-\lambda-h\,\sin 2\theta)^2 + h^2\cos^2 2\theta \\
&= (x-\lambda)^2 + h^2 - 2\,h\,\sin 2\theta(x-\lambda) \\
&\ge (x-\lambda)^2 + h^2 - \frac{6}{5}\,|x-\lambda|\,h.
\end{aligned}$$

Here we have used $|\sin 2\theta| \le \frac{3}{5}$, which follows from $|\tan\theta| \le \frac{1}{3}$. Setting $u(x) = \dfrac{x-\lambda}{h(x)}$ we deduce

$$\begin{aligned}
K_1^2 &\le \int_a^b \frac{(\lambda-x)h' + h}{(x-\lambda)^2 + h^2 - \frac{6}{5}\,|x-\lambda|\,h}\,dx = \int_a^b \frac{u'}{u^2+1-\frac{6}{5}\,|u|}\,dx \\
&\le 2\int_0^\infty \frac{du}{u^2+1-\frac{6}{5}\,|u|} = \frac{5\pi}{4} + \frac{5}{2}\arctan\frac{3}{4} \le 1.763\,\pi.
\end{aligned}$$

b) We will now bound $\partial_t\Psi(t,x)\,u$ and note that

$$\begin{aligned}
\partial_t\Psi(t,x) &= -i\,\Gamma(B,x)^{1/2}(x-i\eta-i\,h\,e^{-2i\theta}-B+i\,t\,C)^{-1}C(x-i\eta-i\,h\,e^{-2i\theta}-B+i\,t\,C)^{-1} \\
&= -i\,\Psi(t,x)\,C\,\Gamma(B,x)^{-1/2}\,\Psi(t,x).
\end{aligned}$$

Using Lemma 11 we get

$$\begin{aligned}
\|\partial_t\Psi(t,x)\,u\| &= \sup_{\|v\|\le 1} \big|\langle \partial_t\Psi(t,x)\,u, v\rangle\big| \\
&\le \sup_{\|v\|\le 1} \big|\langle C\,\Gamma(B,x)^{-1/2}\Psi(t,x)u, \Psi(t,x)^*v\rangle\big| \\
&\le 2\,\|\Psi(t,x)\,u\|\,K_2(t),
\end{aligned} \tag{15}$$

where

$$\begin{aligned}
K_2(t) &= \sup_{\|v\|\le 1} \|\Gamma(B,x)^{1/2}\Psi(t,x)^*v\| \\
&= \sup_{\|v\|\le 1} \|\Gamma(B,x)^{1/2}(x+i\eta+i\,h\,e^{2i\theta}-B-i\,t\,C)^{-1}\Gamma(B,x)^{1/2}v\|.
\end{aligned}$$

Let us consider

$$w := \Psi(t,x)^*v = (x+i\eta+i\,h\,e^{2i\theta}-B-i\,t\,C)^{-1}\Gamma(B,x)^{1/2}v$$

and

$$\Delta := \langle (x+i\eta+i\,h\,e^{2i\theta}-B-i\,t\,C)\,w, w\rangle = \langle v, \Gamma(B,x)^{1/2}w\rangle.$$

We deduce from the Cauchy-Schwarz inequality

$$|\Delta| \le \|\Gamma(B,x)^{1/2}w\|\;\|v\|. \tag{16}$$

We also have

$$|\Delta|^2 = \langle (x-B-\sin 2\theta\, h)\, w, w\rangle^2 + \langle (\eta+\cos 2\theta\, h - t\, C)\, w, w\rangle^2 = |\Delta_1|^2 + |\Delta_2|^2.$$

Using that $\langle C\, w, w\rangle \leq \langle (\eta + \eta'(B-x))w, w\rangle$ and $\eta' = \tan\theta$, we get

$$\Delta_2 = \langle (\eta+\cos 2\theta\, h - t\, C)\, w, w\rangle \geq \langle (\cos 2\theta\, h - \tan\theta\,(B-x))\, w, w\rangle.$$

We introduce

$$X_1 := \langle (B-x)\, w, w\rangle \quad \text{and} \quad X_2 := h\, \langle w, w\rangle.$$

We then have $X_2 \geq 0$ and

$$|\Delta_1| = |X_1 + \sin 2\theta\, X_2|, \quad \Delta_2 \geq \cos 2\theta\, X_2 - \tan\theta\, X_1, \quad \|\Gamma(B,x)^{1/2} w\|^2 = h'\, X_1 + X_2.$$

Assume first $\cos 2\theta\, X_2 - \tan\theta\, X_1 \geq 0$. We then have

$$\begin{aligned} |\Delta|^2 &\geq |X_1 + \sin 2\theta\, X_2|^2 + (\cos 2\theta\, X_2 - \tan\theta\, X_1)^2 \\ &\geq (1+\tan^2\theta) X_1^2 + 2\tan\theta\, X_1\, X_2 + X_2^2. \end{aligned}$$

Therefore

$$\begin{aligned} 10\,|\Delta|^2 - 9\,\|\Gamma(B,x)^{1/2} w\|^4 &\geq \big(10(1+\tan^2\theta) - 9\, {h'}^2\big)\, X_1^2 + 2(10\, \tan\theta - 9h')\, X_1 X_2 + X_2^2 \\ &\geq (X_2 + (10\tan\theta - 9h')X_1)^2 + 10\,(1 - 9(\tan\theta - h')^2) X_1^2. \end{aligned}$$

Using that $|\tan\theta - h'| \leq 1/3$ we find

$$\|\Gamma(B,x)^{1/2} w\|^4 \leq \frac{10}{9} |\Delta|^2 \leq \frac{10}{9}\, \|\Gamma(B,x)^{1/2} w\|^2\ \|v\|^2,$$

which implies $\|\Gamma(B,x)^{1/2} w\| \leq \frac{\sqrt{10}}{3}\, \|v\|$.

Consider now the remaining case, $\tan\theta\, X_1 > \cos 2\theta\, X_2$. We then have

$$\begin{aligned} \frac{\|\Gamma(B,x)^{1/2} w\|^2}{\Delta_1} &= \frac{|h' X_1 + X_2|}{|X_1 + \sin 2\theta\, X_2|} = \frac{|h' X_1 \tan\theta + X_2 \tan\theta|}{|X_1 \tan\theta + 2\, \sin^2\theta\, X_2|} \\ &\leq \max(|h'|, \frac{|h' \cos 2\theta + \tan\theta|}{|cos 2\theta + 2\sin^2\theta|}) = \max(|h'|, |h' \cos 2\theta + \tan\theta|) \leq 1. \end{aligned}$$

This shows $\|\Gamma(B,x)^{1/2} w\| \leq |\Delta|$ and consequently

$$\|\Gamma(B,x)^{1/2} w\| \leq \|v\| \leq \frac{\sqrt{10}}{3}\, \|v\|.$$

Returning to K_2 we hence deduce

$$K_2(t) = \sup_{\|v\|\leq 1} \|\Gamma(B,x)^{1/2} \Psi(t,x)^* v\| = \sup_{\|v\|\leq 1} \|\Gamma(B,x)^{1/2} w\| \leq \frac{\sqrt{10}}{3},$$

and from (15)

$$\|\partial_t \Psi(t,x)\, u\| \leq 2\, \frac{\sqrt{10}}{3}\, \|\Psi(t,x)\, u\|.$$

By integration of this differential inequality we conclude

$$\|\Psi(1,x)\, u\| \leq \exp(2\, \frac{\sqrt{10}}{3})\, \|\Psi(0,x)\, u\| \leq 10.93\sqrt{\pi}\, \|u\|.$$

□

We are now able to bound the last two terms in (9).

Lemma 13. *We have the estimate*

$$\|G_+P(A^*) - G_+P(B)\| \leq 14.74.$$

Proof. As in the previous lemma we omit the subscript +. With $u \in H^n$, $v \in H^m$, we start from

$$\begin{aligned} 2\pi\|GP(A^*) - GP(B)\| &= \\ &\sup_{\|u\|\leq 1, \|v\|\leq 1} \Big| \int_a^b e^{-2i\theta} \langle P(\tau) \big((\bar{\tau} - ih(1+e^{-2i\theta}) - A^*)^{-1} \\ &\qquad - (\bar{\tau} - ih(1+e^{-2i\theta}) - B)^{-1} \big) u, v \rangle \, d\tau \Big| \\ &\leq \sup_{\|u\|\leq 1, \|v\|\leq 1} \int_a^b \Big| \langle C\, P(\tau) \big((\bar{\tau} - ih(1+e^{-2i\theta}) - A^*)^{-1} u, (\tau + ih(1+e^{2i\theta}) - B)^{-1} v \rangle \Big| \, |d\tau|. \end{aligned}$$

We remark that $|d\tau|/dx = |\tau'(x)| \leq \sqrt{37}/6$. Recalling the definition (14) we obtain, using $\|P(\tau)\| \leq 1$ and Lemma 11,

$$2\pi\|GP(A^*) - GP(B)\| \leq 2\,\frac{\sqrt{37}}{6} \sup_{\|u\|\leq 1, \|v\|\leq 1} \int_a^b \|\Psi(1, x)\, u\| \; \|\Psi(0, x)^*\, v\| \, dx.$$

In view of Lemma 12 this yields

$$2\pi\|GP(A^*) - GP(B)\| \leq 2\,\frac{\sqrt{37}}{6}\, \pi\, 1.33 \times 10.93 < 2\,\pi \times 14.74.$$

□

With a similar proof we find

$$\|G_-P(A^*) - G_-P(B)\| \leq 14.74.$$

This inequality combined with Lemma 10 and Corollary 8 completes the proof of Theorem 1.

6 Appendix: an inequality for plane convex domains

We set $B(\omega, r) = \{z \in \mathbb{C}\,; |z - \omega| < r\}$.

Lemma 14. *Let Ω be a convex domain in the complex plane satisfying $B(\omega, r) \subset \Omega \subset B(\omega, R)$, and let $R \geq 2r$. Then we have*

$$\int_{\sigma \in \partial\Omega} \big| d \log |\sigma - \omega| \,\big| \leq 2 \log \frac{R^4}{r^2(R^2 - 2r^2)}. \tag{17}$$

This estimate is best possible.

Proof. It is sufficient to consider a polygonal convex domain Ω. Using a counterclockwise orientation of the boundary, we denote by $\sigma_0, \sigma_{\frac{1}{2}}, \sigma_1, \ldots, \sigma_{n+\frac{1}{2}}, \sigma_{n+1} = \sigma_0$, the points on $\partial\Omega$ corresponding to the successive local maxima "$|\sigma_j - \omega|$" and local minima "$|\sigma_{j+\frac{1}{2}} - \omega|$" of $|\sigma - \omega|$. Note that $n \geq 2$. We then have

$$\begin{aligned} \Phi(\Omega) := \int_{\partial\Omega} \big|d \log|\sigma - \omega|\,\big| \;&=\; \sum_{j=0}^{n} \Big(\log \frac{|\sigma_j - \omega|}{|\sigma_{j+\frac{1}{2}} - \omega|} + \log \frac{|\sigma_{j+1} - \omega|}{|\sigma_{j+\frac{1}{2}} - \omega|} \Big) \\ &=\; |\log J(\alpha)| + |\log J(\beta)|, \end{aligned}$$

where we have set

$$\alpha_j = \arccos \frac{|\sigma_{j+\frac{1}{2}} - \omega|}{|\sigma_j - \omega|}, \quad \beta_j = \arccos \frac{|\sigma_{j+\frac{1}{2}} - \omega|}{|\sigma_{j+1} - \omega|}, \quad J(\alpha) = \prod_{j=0}^{n} \cos \alpha_j.$$

We remark that, with $\gamma = \arccos \dfrac{r}{R}$,

$$\begin{aligned} &J(\alpha) = J(\beta), \qquad 0 < \alpha_j \leq \gamma, \qquad 0 < \beta_j \leq \gamma, \qquad \gamma \in (\frac{\pi}{3}, \frac{\pi}{2}), \\ &\qquad 0 < \alpha_j + \beta_j \leq \arg \frac{\sigma_{j+1} - \omega}{\sigma_j - \omega}, \qquad \sum_{j=0}^{n} (\alpha_j + \beta_j) \leq 2\pi. \end{aligned}$$

Since $\Phi(\Omega) = -2 \log J(\alpha)$, the inequality (17) is equivalent to

$$J(\alpha) \geq \cos^2\gamma \; \cos(\pi - 2\gamma) = \frac{r^2(R^2 - 2r^2)}{R^4}. \tag{18}$$

Note that, with $\alpha_0 = \alpha_1 = \beta_0 = \beta_1 = \gamma$, $\alpha_2 = \beta_2 = \pi - 2\gamma$, $n = 2$, equality is attained in (18), and thus also in (17), which shows the "best possible part" of the lemma.

We can assume that $\sum_{j=0}^{n} \alpha_j \leq \pi$ (otherwise we exchange α and β). We then set

$$J_0 = \min\{J(\alpha)\,; \sum_{j=0}^{n} \alpha_j \leq \pi,\ 0 \leq \alpha_j \leq \gamma,\ \text{for } j = 0, 1, \ldots, n\}.$$

It is easily seen that the minimum is attained and, since $n \geq 2$,

$$J_0 = \min\{J(\alpha)\,; \sum_{j=0}^{n} \alpha_j = \pi,\ 0 \leq \alpha_j \leq \gamma,\ \text{for } j = 0, 1, \ldots, n\}.$$

From the optimality conditions we have, at a minimizing point α,

$$-J_0 \tan \alpha_j + \lambda = 0, \quad \text{for all } j \text{ such that } 0 < \alpha_j < \gamma,$$

where λ is the Lagrange multiplier related to the constraint $\sum_j \alpha_j = \pi$. We deduce that

$$\text{either } \alpha_j = 0, \quad \text{or } \alpha_j = \arctan \frac{\lambda}{J_0} \ (\ell \text{ times}), \quad \text{or } \alpha_j = \gamma \ (k \text{ times}).$$

We consider the only three different possibilities for k :

- $k = 0$, then $l \geq 3$ and $J(\alpha) = \big(\cos \dfrac{\pi}{\ell}\big)^{\ell} \geq \big(\cos \dfrac{\pi}{3}\big)^3 = \dfrac{1}{8}$.
- $k = 1$, then $l \geq 2$ and $J(\alpha) = \cos\gamma \, \big(\cos \dfrac{\pi - \gamma}{\ell}\big)^{\ell} \geq \cos\gamma \, \big(\cos \dfrac{\pi - \gamma}{2}\big)^2$.

- $k = 2$, then $l \geq 1$ and $J(\alpha) = \cos^2\gamma\,\big(\cos\dfrac{\pi-2\gamma}{\ell}\big)^\ell \geq \cos^2\gamma\,\cos(\pi-2\gamma)$.

Hence

$$J(\alpha) \geq \min\big(\frac{1}{8}, \cos\gamma\,\big(\cos\frac{\pi-\gamma}{2}\big)^2, \cos^2\gamma\,\cos(\pi-2\gamma)\big) = \cos^2\gamma\,\cos(\pi-2\gamma).$$

This shows (18) and thus completes the proof. □

7 Some concluding remarks

Remark 1. The estimate $\mathcal{Q}_{cb} \leq 56$ obtained in this paper is not optimal. There is no doubt that refinements are possible in Lemmas 9, 12, 13, and especially in Lemma 10, which would decrease this bound. We know how to replace 56 by 34 but we have favoured simplicity in the proofs. We are convinced that our estimate is very pessimistic, but to improve it drastically (recall that our conjecture is $\mathcal{Q}_{cb} = 2$), it is clear that we have to find a completely different method. In the simple case when Ω is a disk, the constant 2 in the inequality (5) has been obtained using a sophisticated representation formula due to Ando (cf. [2]).

Remark 2. We have shown that the map u_A from the algebra of rational functions bounded on $W(A)$ into the C^*-algebra $\mathcal{L}(H)$, defined by $u_A(r) = r(A)$, is completely bounded with a complete norm $\|u_A\|_{cb} \leq \mathcal{Q}_{cb}$. A direct application of Paulsen's theorem, see [4], gives:

There exists an invertible operator $S \in \mathcal{L}(H)$ such that $\|S\|\,\|S^{-1}\| \leq \mathcal{Q}_{cb}$ and

$$\|r(S^{-1}AS)\| \leq \sup_{z\in W(A)} |r(z)|, \qquad \textit{for any rational function } r \textit{ bounded in } W(A),$$

or, equivalently, $W(A)$ is a *spectral set* for the operator $S^{-1}AS$. Note that we have a reduction of the numerical range: $W(S^{-1}AS) \subset W(A)$.

Remark 3. We also deduce from a result due to Arveson (see [1], [4]) that there exist a larger Hilbert space $\mathcal{K}$, containing H as a subspace, and a normal operator N acting on $\mathcal{K}$, with spectrum $\sigma(N) \subset \partial W(A)$, such that

$$r(A) = S\,P_H\,r(N)|_H\,S^{-1}, \qquad \text{for any rational function } r \text{ bounded in } W(A),$$

where P_H denotes the orthogonal projection from $\mathcal{K}$ onto H. In other words A is similar to an operator having a *normal $\partial W(A)-$dilation*.

References

[1] W. Arveson : Subalgebras of C^*-algebras. II, *Acta Math.* **128** (1972), 271-308.

[2] M. Crouzeix, Bounds for analytic functions of matrices, *Int. Equ. Op. Th.* **48** (2004), 461–477.

[3] M. Crouzeix, Operators with numerical range in a parabola, *Arch. Math.*, **82** (2004), 517-527.

[4] V. Paulsen, *Completely bounded maps and operator algebras*, Cambridge Univ. Press, 2002.

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