Convex domains and K-spectral sets

Catalin Badea, Michel Crouzeix and Bernard Delyon

Abstract

Let Ω be an open convex domain of $\mathbb{C}$. We study constants K such that Ω is K-spectral or complete K-spectral for each continuous linear Hilbert space operator with numerical range included in Ω. Several approaches are discussed.

2000 Mathematical subject classifications : 47A12 ; 47A25

1 Introduction

Let H be a complex Hilbert space and let $\mathcal{L}(H)$ denote the C*-algebra of all continuous linear operators on H. For $A \in \mathcal{L}(H)$ its numerical range $W(A)$ is defined by

$$W(A) = \{\langle Ax, x\rangle\,;\, x \in H,\ \|x\| = 1\}.$$

Here $\langle x, y\rangle$ is the inner product in H and $\|x\| = \langle x, x\rangle^{1/2}$ the corresponding norm. Recall that $W(A)$ is a convex subset of $\mathbb{C}$ (the Toeplitz-Hausdorff theorem) and that the spectrum of A is contained in $\overline{W(A)}$. If a rational function r is bounded on the numerical range, then it has no pole in $\overline{W(A)}$; consequently $r(A)$ is well-defined and belongs to $\mathcal{L}(H)$.

The aim of this paper is to study the following constant

$$C(\Omega) := \sup\{\|r(A)\|\,;\, \overline{W(A)} \subset \Omega, |r(z)| \le 1, \forall z \in \Omega\}, \tag{1.1}$$

where Ω is an open convex subset of the complex plane ($\emptyset \ne \Omega \ne \mathbb{C}$). In this definition the supremum is taken over all complex Hilbert spaces H, all continuous linear operators A on H and all rational functions $r : \mathbb{C} \to \mathbb{C}$ satisfying the prescribed constraints. Recall also that the operator norm $\|r(A)\|$ is defined by

$$\|r(A)\| = \sup\{\|r(A))x\|\,;\, x \in H, \|x\| \le 1\}.$$

It is also interesting to consider the completely bounded analogue of $C(\Omega)$ defined by

$$C_{cb}(\Omega) := \sup\{\|R(A)\|\,;\, \overline{W(A)} \subset \Omega, \|R(z)\| \le 1, \forall z \in \Omega\}. \tag{1.2}$$

As in (1.1), the supremum is taken over all complex Hilbert spaces H, all continuous linear operators A on H (satisfying the prescribed constraints), but now R runs among the rational functions with matrix values of size $m \times n$ ($R : \mathbb{C} \to \mathbb{C}^{m,n}$) and the supremum is taken also over all m and n.

We clearly have $1 \le C(\Omega) \le C_{cb}(\Omega)$ and these constants only depend on the shape of Ω; indeed, they are invariant under similarities and symmetries. We also have

$$C(\Omega) = \sup\{\|r(A)\|\,;\, W(A) \subset \overline{\Omega}, |r(z)| \le 1, \forall z \in \Omega\},$$

and a similar formula holds for $C_{cb}(\Omega)$.

Recall that Ω is said to be a *K-spectral set* [22] (resp. a *spectral set*) for an operator A if the inequality

$$\|r(A)\| \le K \sup_{z\in\Omega} |r(z)|, \qquad (\text{resp. } \|r(A)\| \le \sup_{z\in\Omega} |r(z)|)$$

holds for all rational functions without pole in the spectrum of A. Therefore our definition (1.1) means that Ω *is a* $C(\Omega)-$*spectral set* for all operators A such that $\overline{W(A)} \subset \Omega$. Similarly (1.2) means that Ω *is a complete* $C_{cb}(\Omega)-$*spectral set* for these operators.

A famous result, due to J. von Neumann [26], states the following :
if Π *is a half-plane, then* Π *is a spectral set for A, for any operator A with* $W(A) \subset \Pi$*;*
in other words $C(\Pi) = 1$. We also have $C_{cb}(\Pi) = 1$. These are the first known estimates for the constants $C(\Omega)$ and $C_{cb}(\Omega)$. Since then it has been shown that

- $C(\Omega) < +\infty$, for all bounded open convex subsets Ω, see [11],
- $C_{cb}(S) < 2 + 2/\sqrt{3}$, for all convex sectors or all strips S, see [9],
- $C(D) = 2$, for all disks D, see [6, 20],
- $C_{cb}(P) < 4.75$, for all parabolic domains P, see [7],
- $C_{cb}(\Omega) < 57$, for all open convex subsets Ω, see [8].

Excepting $C_{cb}(\Pi) = 1$ and $C(D) = 2$, these estimates are far from being optimal ; the conjecture $\sup_\Omega C(\Omega) = 2$ has been proposed in [6], and then the conjecture $\sup_\Omega C_{cb}(\Omega) = 2$ in [8].

The boundedness of the constants $C(\Omega)$ and $C_{cb}(\Omega)$ allows to extend the rational functional calculus for operators A satisfying $W(A) \subset \overline{\Omega}$ to more general (holomorphic) functions. Furthermore, if Ω is unbounded, and after adding a technical resolvent condition, a suitable functional calculus can be constructed for unbounded operators ; we refer for that to [14]. Up to now, the boundedness of these constants has allowed to obtain some new results : a proof of the Burkholder conjecture in probability theory [11], a shorter proof of the Boyadzhiev-de Laubenfels theorem (concerning decomposition for group generators, see [14]), a characterization for generators of cosine functions [2, 7, 14], and a characterization of similarities of ω-accretive operators [18].

Let us mention some other consequences. Assuming $W(A) \subset \overline{\Omega}$, the inequality

$$\|R(A)\| \le C_{cb}(\Omega) \sup_{z\in\Omega} \|R(z)\|,$$

which holds for all rational functions with matrix values, means that the homomorphism u_A from the algebra of rational functions bounded on Ω into the C^*-algebra $\mathcal{L}(H)$, defined by $u_A(r) = r(A)$, is completely bounded with $\|u_A\|_{cb} \le C_{cb}(\Omega)$. A direct application of Paulsen's Theorem (see [21] or [22, Theorem 9.1]) gives

There exists an invertible operator $S \in \mathcal{L}(H)$*, with* $\|S\|\,\|S^{-1}\| \le C_{cb}(\Omega)$*, such that the domain* Ω *is a complete spectral set for* $S\,A\,S^{-1}$*.*

We deduce then from a result due to Arveson (see [3] or [22, Corollary 7.7]) that there exists a larger Hilbert space K containing H as a subspace (with the same inner product) and a normal operator N acting on K, with spectrum $\sigma(N) \subset \partial\Omega$, such that, for all rational functions r bounded in Ω, we have

$$r(A) = S^{-1}\, P_H\, r(N)|_H\, S.$$

Here P_H denotes the orthogonal projection from K onto H. In other words, if $W(A) \subset \overline{\Omega}$, then A is similar to an operator having a *normal* $\partial\Omega-$*dilation*. We would like to stress here that our methods give sharp estimates for the similarity constant $\|S\|\,\|S^{-1}\|$. In particular, we obtain a similarity constant which is independent of Ω.

Another motivation for our study is that estimates for $C(\Omega)$ and $C_{cb}(\Omega)$ have an interesting potential of applications in numerical analysis. In computational linear algebra for instance, the popular Krylov type methods for solving large linear systems $Ax = b$ are based on polynomial approximations of A^{-1}. We refer to [4], where the authors are using some results of the present paper to improve known error estimates for the GMRES method. Also, the time discretizations of parabolic type P.D.E. use rational approximations of the exponential. If the boundedness is often sufficient for theoretical developments, sufficiently good estimates of our constants are desirable for numerical applications.

The goal of this paper is to present different approaches which can be used for estimating $C(\Omega)$ and $C_{cb}(\Omega)$; the outline is as follows. The first sections are based on appropriate integral representations of $r(A)$ or $R(A)$; the positivity (for convex domains) of the double layer potential plays an important role. In Section 2, for a bounded convex domain we show that

$$C_{cb}(\Omega) \leq 2 + \pi + \inf_{\omega \in \Omega} \mathrm{TV}(\log|\sigma - \omega|). \tag{1.3}$$

Here $\mathrm{TV}(\log|\sigma-\omega|)$ denotes the total variation of $\log(|\sigma - \omega|)$ as σ runs around $\partial\Omega$. In the unbounded case we obtain in Section 3 the inequality

$$C_{cb}(\Omega) \leq 1 + \frac{2}{\pi}\int_{\alpha}^{\pi/2} \frac{\pi - x + \sin x}{\sin x}\, dx, \tag{1.4}$$

if Ω contains a sector of positive angle 2α, $0 < \alpha \leq \frac{\pi}{2}$. Another representation, based on the solution of the C. Neumann problem for the double layer potential, is developped in Section 4. Connexions with dilation theorems are indicated. Section 5 is devoted to the similarity approach ; it gives a complete answer for the disk case $C(D) = C_{cb}(D) = 2$ and it is used to show that $\sup_\Omega C_{cb}(\Omega, 2) = 2$, where $C_{cb}(\Omega, 2)$ is defined similarly as $C_{cb}(\Omega)$ but the supremum is taken now only over the 2×2 matrices.

2 The case of a bounded convex

Let Ω be a proper open convex of the complex plane. On the counterclockwise oriented boundary $\partial\Omega$, we consider the generic point σ of arclength s ; we denote by $\nu = \frac{1}{i}\frac{d\sigma}{ds}$ the unit outward normal (which exists a.e.). Let $A \in \mathcal{L}(H)$ be an operator. We introduce the function $\mu(\sigma, z)$, the half-plane $\Pi_\sigma \supset \Omega$ and the self-adoint operator $\mu(\sigma, A)$ defined by

$$\begin{aligned}
\mu(\sigma, z) &= \frac{1}{\pi}\frac{d}{ds}\big(\arg(\sigma - z)\big) = \frac{1}{2\pi}\Big(\frac{\nu}{\sigma - z} + \frac{\bar\nu}{\bar\sigma - \bar z}\Big), \\
\Pi_\sigma &= \{z\,;\mu(\sigma, z) > 0\} = \{z\,;\mathrm{Re}\,\bar\nu(\sigma - z) > 0\}, \\
\mu(\sigma, A) &= \tfrac{1}{2\pi}\big(\nu(\sigma - A)^{-1} + \bar\nu(\bar\sigma - A^*)^{-1}\big) \quad \text{(if } \sigma \text{ belongs to the resolvent of } A).
\end{aligned}$$

Lemma 2.1. *We assume that the convex domain Ω contains the spectrum of $A \in \mathcal{L}(H)$. Then the condition $W(A) \subset \Omega$ is equivalent to $\mu(\sigma, A) > 0,\ \forall \sigma \in \partial\Omega$. When this condition is satisfied, and if Ω is bounded and g is a continuous function bounded by 1 on $\partial\Omega$, then we have*

$$\Big\|\int_{\partial\Omega} g(\sigma)\,\mu(\sigma, A)\,ds\Big\| \leq 2.$$

Proof. We have (setting $w=(\sigma-A)^{-1}v$)

$$\begin{aligned}\mu(\sigma,A)>0 \quad &\Longleftrightarrow \quad \forall\, 0\neq v\in H, \quad \text{Re}\ \bar\nu\langle(\bar\sigma-A^*)^{-1}v,v\rangle>0\\ &\Longleftrightarrow \quad \forall\, 0\neq w\in H, \quad \text{Re}\ \bar\nu\langle(\sigma-A)w,w\rangle>0\\ &\Longleftrightarrow \quad W(A)\subset\Pi_\sigma.\end{aligned}$$

The equivalence follows from the convexity property $\Omega=\cap_{\sigma\in\partial\Omega}\Pi_\sigma$.

We deduce from the Cauchy formula that $\int_{\partial\Omega}\mu(\sigma,A)\,ds=2$. We set $\Gamma=\int_{\partial\Omega}g(\sigma)\,\mu(\sigma,A)\,ds$ and consider $u,v\in H$ such that $\|u\|=\|v\|=1$. From the positivity of $\mu(\sigma,A)$ we have

$$|\langle\mu(\sigma,A)\,u,v\rangle|\le\tfrac12\,\langle\mu(\sigma,A)\,u,u\rangle+\tfrac12\,\langle\mu(\sigma,A)\,v,v\rangle;$$

therefore

$$|\langle\Gamma\,u,v\rangle|\le\tfrac12\int_{\partial\Omega}\langle\mu(\sigma,A)\,u,u\rangle\,ds+\tfrac12\int_{\partial\Omega}\langle\mu(\sigma,A)\,v,v\rangle\,ds=\|u\|^2+\|v\|^2=2.$$

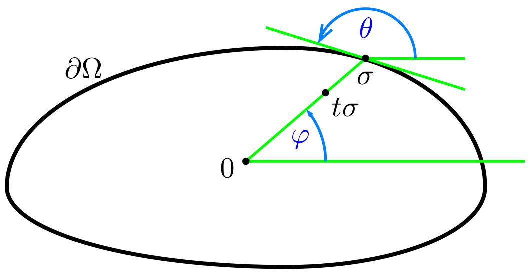
The result follows from $\|\Gamma\|=\sup\{|\langle\Gamma\,u,v\rangle|\,;\|u\|=\|v\|=1\}$. □

Remark. With an easy modification this proof is also valid if g is a matrix valued function.

For the remaining part of this section we assume that the domain Ω is bounded and that the origin 0 belongs to Ω. Now we introduce the angles θ and φ such that

$$e^{i\theta}=\tfrac{d\sigma}{ds}$$

$$\varphi=\arg\sigma$$



We can assume that the determinations of θ and φ are choosed such that $0<\theta-\varphi<\pi$. In order to avoid some technical difficulties, we initially assume that

$$\theta\ \text{ is a } C^1 \text{ function of } s \text{ and } \theta'(s)>0,\forall\, s\in[0,L], \qquad \text{where } L \text{ is the length of } \partial\Omega. \tag{2.1}$$

This assumption allows to consider also s, σ and φ as C^1 functions of θ.

Lemma 2.2. *Let r be a rational function bounded in a domain Ω satisfying (2.1). Then we have*

$$r(z)=\int_{\partial\Omega}r(\sigma)\mu(\sigma,z)\,ds+\int_0^{2\pi}Jr(\sigma,\bar z)\,d\theta, \qquad \forall z\in\Omega, \tag{2.2}$$

where $$Jr(\sigma,\bar z):=\frac{1}{\pi}\int_0^1\frac{e^{2i\theta}\sigma(\bar\sigma-\bar z)\,r(t\sigma)}{\big((t-1)\sigma+e^{2i\theta}(\bar\sigma-\bar z)\big)^2}\,dt.$$

Furthermore, if $|r|\le1$ in Ω, then we have the estimate

$$|Jr(\sigma,\bar z)|\le\tfrac12+|\cot(\theta-\varphi)|, \qquad \forall\, z\in\Pi_\sigma. \tag{2.3}$$

Proof. a) For a given point $z \in \Omega$ we introduce (σ and θ being functions of s)

$$\begin{aligned} u(t,s) &:= (t-1)\sigma + e^{2i\theta}(\bar\sigma - \bar z), \\ v(t,s) &:= \frac{\sigma\, r(t\sigma)}{u}, \quad w(t,s) := \frac{t\, r(t\sigma)}{u}\,\frac{d\sigma}{ds}. \end{aligned}$$

We first notice that

$$\frac{\partial}{\partial s}(u\,v) - \frac{\partial}{\partial t}(u\,w) = 0 \quad \text{and} \quad d\sigma = e^{2i\theta} d\bar\sigma;$$

therefore

$$u\,\Big(\frac{\partial w}{\partial t} - \frac{\partial v}{\partial s}\Big) = v\,\frac{\partial u}{\partial s} - w\,\frac{\partial u}{\partial t} = \frac{r(t\sigma)}{u}\,2\,i\,e^{i\theta}\,\sigma\,(\bar\sigma - \bar z)\frac{d\theta}{ds},$$

$$\text{and} \qquad \frac{\partial w}{\partial t} - \frac{\partial v}{\partial s} = 2\,i\,\frac{e^{2i\theta}\sigma(\bar\sigma-\bar z)\,r(t\sigma)}{\big((t-1)\sigma + e^{2i\theta}(\bar\sigma-\bar z)\big)^2}\,\frac{d\theta}{ds}.$$

The function $u(t,s)$ does not vanish for $(t,s) \in [0,1]\times[0,L]$; this shows that the functions v, $\frac{\partial v}{\partial s}$, w, $\frac{\partial w}{\partial t}$ are well defined and continuous on this set. Therefore we can write

$$\begin{aligned} 2\,\pi\,i \int_0^{2\pi} Jr(\sigma,\bar z)\,d\theta &= \int_0^L\!\!\int_0^1 \Big(\frac{\partial w}{\partial t} - \frac{\partial v}{\partial s}\Big)\,dt\,ds \\ &= \int_0^L\!\!\int_0^1 \frac{\partial w}{\partial t}\,dt\,ds - \int_0^1\!\!\int_0^L \frac{\partial v}{\partial s}\,ds\,dt \\ &= \int_0^L w(1,s)\,ds = \int_{\partial\Omega} \frac{r(\sigma)}{\bar\sigma - \bar z}\,d\bar\sigma. \end{aligned}$$

We deduce that

$$\int_{\partial\Omega} r(\sigma)\mu(\sigma,z)\,ds + \int_0^{2\pi} Jr(\sigma,\bar z)\,d\theta = \frac{1}{2\pi i}\int_{\partial\Omega}\frac{r(\sigma)}{\sigma - z}\,d\sigma,$$

and then (2.2) follows from the Cauchy formula.

b) We remark that $Jr(\sigma,\bar z)$ is antiholomorphic for $z \in \Pi_\sigma$, bounded and continuous in $\bar\Pi_\sigma$ and vanishing as $z \to \infty$. Using the maximum principle we obtain for $z \in \Pi_\sigma$

$$\begin{aligned} |Jr(\sigma,\bar z)| &\le \sup_{\zeta\in\partial\Pi_\sigma} |Jr(\sigma,\bar\zeta)| = \sup_{x\in\mathbb{R}} |Jr(\sigma, \bar\sigma + x e^{-i\theta})| \\ &\le \sup_{x\in\mathbb{R}} \int_0^1 \frac{|\sigma|\,|x|}{\pi\,|(t-1)\sigma - x e^{i\theta}|^2}\,dt \\ &\le \max_{\varepsilon=\pm1}\int_0^\infty \frac{d\tau}{\pi\,|\tau e^{i\varphi} - \varepsilon e^{i\theta}|^2} = \frac{\max(\theta-\varphi, \pi-\theta+\varphi)}{\pi\,\sin(\theta-\varphi)}. \end{aligned}$$

We have used above the change of variables $\tau = (t{-}1)|\sigma|/|x|$. Now we remark that the inequality $\max(u, \pi - u) \le \frac{\pi}{2}\sin u + \pi|\cos u|$ holds for any $u = \theta - \varphi \in (0,\pi)$, which shows that

$$|Jr(\sigma,\bar z)| \le \frac{1}{2} + |\cot(\theta-\varphi)|.$$

□

Theorem 2.3. *For all bounded convex domains Ω we have the estimate*

$$C(\Omega) \le C_{cb}(\Omega) \le 2 + \pi + \phi_\Omega, \qquad \textit{where } \phi_\Omega := \inf_{\omega\in\Omega} TV(\log|\sigma - \omega|).$$

Proof. Without loss of generality we can assume that $\omega = 0$ is the origin of the complex plane. We assume initially that (2.1) is satisfied. Let A be an operator with $\overline{W(A)} \subset \Omega$ and r a rational function bounded by 1 in Ω. We can replace z by A and $\bar{z}$ by A^* in the proof of the previous lemma. We obtain

$$2\,\pi\,i \int_0^{2\pi} Jr(\sigma, A^*)\,d\theta = \int_{\partial\Omega} r(\sigma)\,(\bar{\sigma} - A^*)^{-1} d\bar{\sigma},$$

and then, from the Cauchy formula,

$$r(A) = \int_{\partial\Omega} r(\sigma)\mu(\sigma, A)\,ds + \int_0^{2\pi} Jr(\sigma, A^*)\,d\theta.$$

¿From Lemma 2.1 we know that $\| \int_{\partial\Omega} r(\sigma)\mu(\sigma, A)\,ds\| \leq 2$. Using the von Neumann inequality for the half-plane Π_σ we have

$$\|Jr(\sigma, A^*)\| \leq \sup_{z\in\Pi_\sigma} |Jr(\sigma, \bar{z})| \leq \frac{1}{2} + |\cot(\theta - \varphi)|.$$

Therefore

$$\|r(A)\| \leq 2 + \pi + \int_0^{2\pi} |\cot(\theta - \varphi)|\,d\theta = 2 + \pi + \int_0^{2\pi} |\cot(\theta - \varphi)|\,d\varphi;$$

indeed $\int_0^{2\pi} |\cot(\theta - \varphi)|\,d(\theta - \varphi) = 0$. Setting $\rho(\varphi) := |\sigma|$, we have $\cot(\theta - \varphi) = \rho'(\varphi)/\rho(\varphi)$ and thus

$$\|r(A)\| \leq 2 + \pi + \int_0^{2\pi} \Big|\frac{\rho'(\varphi)}{\rho(\varphi)}\Big|\,d\varphi = 2 + \pi + TV(\log\rho).$$

We obtain the estimate $C(\Omega) \leq 2 + \pi + \phi_\Omega$.

Now we remark that the von Neumann inequality, as well as Lemma 2.1, are also valid for rational functions with matrix values ; therefore we obtain in the same way $C_{cb}(\Omega) \leq 2 + \pi + \phi_\Omega$. The use of Lemma 2.4 hereafter (with $K = \overline{W(A)}$) allows to extend the results to all bounded domains Ω. □

Remark 2.1. For instance, if Ω is an ellipse with major axis $2a$ and minor axis $2b$, then we have $\phi_\Omega = 4\log\frac{a}{b}$; $\phi_\Omega = 0$ if Ω is a disk. For some domains Ω it is dificult to compute exactly ϕ_Ω. For these domains, the following geometrical estimate given in [8] can be useful

$$\phi_\Omega \leq 2\,\log\frac{\tau_\Omega^4}{\tau_\Omega^2 - 2}, \qquad \text{where} \quad \tau_\Omega := \min_{\omega\in\Omega} \frac{\max\{|\sigma - \omega|\,;\sigma \in \partial\Omega\}}{\min\{|\sigma - \omega|\,;\sigma \in \partial\Omega\}}.$$

The quantity τ_Ω can be considered as a rate of flatness of Ω.

If $\tau_\Omega \leq 400\,000$, the above estimate is better than the general bound 57 given in [8].

Lemma 2.4. *We assume that K is a compact subset of a bounded convex domain Ω and that $0 \in \Omega$. Then, for all $\varepsilon > 0$, there exists a bounded convex domain Ω' satisfying (2.1) such that*

$$K \subset \Omega' \subset \Omega, \quad \textit{and} \quad TV_{\partial\Omega'}(\log|\sigma|) \leq TV_{\partial\Omega}(\log|\sigma|) + \varepsilon.$$

Proof. In this lemma we use the notation $r(\varphi) = |\sigma|^{-1} = 1/\rho(\varphi)$, where φ is the argument of the boundary point σ. The convexity of Ω is equivalent to

r is a continuous 2π-periodic function on $\mathbb{R}$, and $r + r''$ is a positive measure.

By a standard mollifier technique we can find a sequence of 2π-periodic functions $r_n \in C^\infty(\mathbb{R})$ such that

$$r_n \to r \text{ uniformly in } \mathbb{R}, \quad \min_x(r_n(x) + r_n''(x)) \geq 0 \text{ and } \quad r_n' \to r' \text{ in } L^1(0, 2\pi).$$

By adding, if needed, to r_n a positive constant η_n, we can assume that $r_n \geq r$ and $\min_x(r_n(x) + r_n''(x)) > 0$. We set

$$\Omega_n := \{\rho\, e^{i\varphi}\,; 0 \leq \rho < 1/r_n(\varphi), \varphi \in [0, 2\pi]\}.$$

It is clear that Ω_n is open, strictly convex, satisfies (2.1), that $\Omega_n \subset \Omega$ and that, for n large enough, $K \subset \Omega_n$. Furthermore, we have

$$TV_{\partial\Omega_n}(\log|\sigma|) = \int_0^{2\pi} \frac{|r_n'(\varphi)|}{r_n(\varphi)}\, d\varphi \to \int_0^{2\pi} \frac{|r'(\varphi)|}{r(\varphi)}\, d\varphi = TV_{\partial\Omega}(\log|\sigma|), \quad \text{as } n \to \infty.$$

The Lemma follows by choosing $\Omega' = \Omega_n$ with a sufficiently large value of n.

□

3 Unbounded domains

We turn now to the case of an unbounded convex domain $\Omega \neq \mathbb{C}$. We assume in this section that Ω contains a convex sector of angle $2\alpha > 0$. Since the constants are only dependent on the shape of Ω, we can assume, without loss of generality, that the value α is maximal and that

$$S_\alpha := \{z \in \mathbb{C}\,; z \neq 0, |\arg z| < \alpha\} \subset \Omega \subset \{z \in \mathbb{C}\,; \operatorname{Re} z > 0\}.$$

We use the same notations as in the previous section. Let $A \in \mathcal{L}(H)$ be an operator with $\overline{W(A)} \subset \Omega$. We remark that

$$\int_{\partial\Omega} \mu(\sigma, A)\, ds = 2 - 2\,\alpha/\pi.$$

Indeed, if we choose $R > \|A\|$ and set $\Omega_R = \{z \in \Omega\,; |z| < R\}$, $C_R = \partial\Omega_R \setminus (\partial\Omega \cap \partial\Omega_R)$, then we have

$$\int_{\partial\Omega_R} \mu(\sigma, A)\, ds = 2, \qquad \int_{C_R} \mu(\sigma, A)\, ds = \int_{C_R} \frac{1}{\pi R}\, ds + O(\frac{\|A\|}{R}) = \frac{2\,\alpha}{\pi} + O(\frac{1}{R}).$$

Thus

$$\int_{\partial\Omega} \mu(\sigma, A)\, ds = \lim_{R\to\infty} \int_{\partial\Omega\cap\partial\Omega_R} \mu(\sigma, A)\, ds = 2 - 2\,\alpha/\pi.$$

This implies that Lemma 2.1 is still valid for unbounded domains, but now with a better estimate : if g is a continuous function bounded by 1 on $\partial\Omega$, then

$$\Big\| \int_{\partial\Omega} g(\sigma)\, \mu(\sigma, A)\, ds \Big\| \leq 2 - 2\,\alpha/\pi.$$

We have now to modify Lemma 2.2 as follows.

Lemma 3.1. *We assume that $\theta \in C^1(\mathbb{R})$ and satisfies $\theta'(s) > 0$ for all real s. Let r be a rational function bounded in Ω. Then we have*

$$r(z) = \int_{\partial\Omega} r(\sigma)\mu(\sigma, z)\, ds + \int_{\pi+\alpha}^{2\pi-\alpha} Kr(\sigma, \bar{z})\, d\theta, \qquad \forall z \in \Omega, \tag{3.1}$$

where $\quad Kr(\sigma, \bar{z}) := -\dfrac{1}{\pi} \displaystyle\int_0^\infty r(\sigma+t)\, \frac{e^{2i\theta}(\bar{\sigma} - \bar{z})}{(t + e^{2i\theta}(\bar{\sigma} - \bar{z}))^2}\, dt.$

Furthermore, if $|r| \leq 1$ in Ω, then we have the estimate

$$|Kr(\sigma,\bar z)| \leq \frac{\max(\theta-\pi, 2\pi-\theta)}{\pi\,\sin\theta}, \qquad \forall\, z\in\Pi_\sigma. \tag{3.2}$$

Proof. a) We first remark that, if $r = 1$ is constant, then $Kr(\sigma,\bar z) = -1/\pi$ and (3.1) is satisfied. It is thus sufficient to suppose $r(\infty) = 0$ in the proof of (3.1).

We note that the angle θ of the tangent vector in the boundary point $\sigma \in \partial\Omega$ is now running in the interval $(\pi+\alpha, 2\pi-\alpha)$. For a fixed $z \in \Omega$ we set

$$u(t,s) := t + e^{2i\theta}(\bar\sigma - \bar z), \quad v(t,s) := \frac{r(\sigma+t)}{u}, \quad w(t,s) := \frac{r(\sigma+t)}{u}\,\frac{d\sigma}{ds}.$$

We remark that the function u does not vanish in $[0,\infty[\times\mathbb{R}$ and that there exists a constant $c > 0$ such that on this set $|u(t,s)| \geq c\,(1+t+|s|)$. We have

$$\begin{aligned} u^2\,(\frac{\partial w}{\partial t} - \frac{\partial v}{\partial s}) &= r'(\sigma+t)\,\frac{d\sigma}{ds}\,u - r(\sigma+t)\,\frac{d\sigma}{ds}\,\frac{\partial u}{\partial t} - r'(\sigma+t)\,\frac{d\sigma}{ds}\,u + r(\sigma+t)\,\frac{\partial u}{\partial s} \\ &= 2\,i\,r(\sigma+t)\,e^{2i\theta}(\bar\sigma-\bar z)\,\frac{d\theta}{ds}. \end{aligned}$$

Thus

$$2\,\pi\,i\,Kr(\sigma,\bar z)\,\frac{d\theta}{ds} = \int_0^\infty \frac{\partial v}{\partial s}(t,s)\,dt - \int_0^\infty \frac{\partial w}{\partial t}\,dt = \frac{r(\sigma)}{e^{2i\theta}(\bar\sigma-\bar z)}\,\frac{d\sigma}{ds} + \int_0^\infty \frac{\partial v}{\partial s}(t,s)\,dt.$$

Noticing that $|\frac{\partial v}{\partial s}(t,s)| \leq C\,(1+t+|s|)^{-3}$, which justifies the use of Fubini, we have

$$\begin{aligned} \int_{\pi+\alpha}^{2\pi-\alpha} Kr(\sigma,\bar z)\,d\theta &= \int_{\partial\Omega} \frac{r(\sigma)}{\bar\sigma-\bar z}\,d\bar\sigma + \int_{-\infty}^{+\infty}\int_0^\infty \frac{\partial v}{\partial s}(t,s)\,dt\,ds \\ &= \int_{\partial\Omega} \frac{r(\sigma)}{\bar\sigma-\bar z}\,d\bar\sigma + \int_0^\infty\int_{-\infty}^{+\infty} \frac{\partial v}{\partial s}(t,s)\,dt\,ds = \int_{\partial\Omega} \frac{r(\sigma)}{\bar\sigma-\bar z}\,d\bar\sigma. \end{aligned}$$

The relation (3.1) follows now from the Cauchy formula.

b) Using the maximum principle we obtain, for $z \in \Pi_\sigma$,

$$\begin{aligned} |Kr(\sigma,\bar z)| &\leq \sup_{\zeta\in\partial\Pi_\sigma} |Kr(\sigma,\bar\zeta)| = \sup_{x\in\mathbb{R}} |Kr(\sigma,\bar\sigma + xe^{-i\theta})| \\ &\leq \sup_{x\in\mathbb{R}} \int_0^\infty \frac{|x|}{\pi\,|t-xe^{i\theta}|^2}\,dt \\ &\leq \max_{\varepsilon=\pm1} \int_0^\infty \frac{d\tau}{\pi\,|\tau-\varepsilon e^{i\theta}|^2} = \frac{\max(\theta-\pi, 2\pi-\theta)}{\pi\,\sin\theta}. \end{aligned}$$

□

Theorem 3.2. *We assume that the convex domain $\Omega \neq \mathbb{C}$ contains a sector of angle 2α, with $0 < \alpha \leq \pi/2$. Then we have the estimate*

$$C(\Omega) \leq C_{cb}(\Omega) \leq 1 + \frac{2}{\pi}\int_\alpha^{\pi/2} \frac{\pi - x + \sin x}{\sin x}\,dx.$$

Proof. As for Theorem 2.3, it is sufficient to look at the estimate of $C(\Omega)$ and we can assume that $\theta \in C^1(\mathbb{R})$ satisfies $\theta'(s) > 0$ for all real s. Let r be a rational function bounded by 1 in Ω. We deduce from (3.1) that

$$r(A) = \int_{\partial\Omega} r(\sigma)\mu(\sigma,A)\,ds + \int_{\pi+\alpha}^{2\pi-\alpha} Kr(\sigma,A^*)\,d\theta.$$

We have seen that

$$\|\int_{\partial\Omega} r(\sigma)\mu(\sigma,A)\,ds\| \le 2-2\,\alpha/\pi, \quad \text{and} \quad \|Kr(\sigma,A^*)\| \le \frac{\max(\theta-\pi,2\pi-\theta)}{\pi\,\sin\theta}.$$

Alltogether that gives

$$\|r(A)\| \le 1+\frac{2}{\pi}\int_{\alpha}^{\pi/2}\frac{\pi-x+\sin x}{\sin x}\,dx$$

and the theorem follows. ☐

4 A potential-theoretic approach

4.1 The Carl Neumann problem

Let $\Omega \neq \emptyset$ be a bounded convex domain of the complex plane. Given a function r, continuous on $\overline{\Omega}$ and harmonic in Ω, the C. Neumann problem [19] for the double layer potential on $\partial\Omega$ is the following : find a function $g \in C(\partial\Omega)$ such that

$$\forall\, z \in \Omega, \quad r(z) = \frac{1}{2}\int_{\partial\Omega} g(\sigma)\,\mu(\sigma,z)\,ds. \tag{4.1}$$

Taking the limit as the point z tends to the boundary $\partial\Omega$, the problem (4.1) is equivalent to

$$\forall\, z \in \partial\Omega, \quad r(z) = \frac{1}{2}\big(g(z)+\int_{\partial\Omega_z} g(\sigma)\,\mu(\sigma,z)\,ds\big), \quad \text{where} \quad \partial\Omega_z = \partial\Omega\setminus\{z\}.$$

This relation can be written (by considering restrictions to $\partial\Omega$)

$$r = \tfrac{1}{2}(I+P)g, \qquad \text{where} \quad Pg(z) = \frac{1}{\pi}\int_{\partial\Omega_z} g(\sigma)\,d\arg(\sigma-z). \tag{4.2}$$

Clearly P is a linear operator acting from $C(\partial\Omega)$ into itself. A harmonic function is uniquely defined by its restriction on the boundary $\partial\Omega$, and any continuous function on this boundary is the trace of such a function. Therefore the invertibility of the operator $I+P$ in $C(\partial\Omega)$ implies existence and uniqueness for (4.1). It is known that this operator is effectively invertible, since we have assumed Ω bounded and convex (see for instance the monograph [17]).

We will mainly restrict our attention to rational functions, bounded in Ω, and introduce the constant

$$C_N(\Omega) = \sup\{2\,\|(I+P)^{-1}r\|_{L^\infty(\partial\Omega)}\,; r \text{ rational function, } |r| \le 1 \text{ in } \Omega\}.$$

In this definition we assume that r acts from $\mathbb{C}$ into $\mathbb{C}$, but the constant would be unchanged by considering matrix-valued rational functions.

Thus, if R is a matrix-valued rational function satisfying $\|R\| \le 1$ in Ω, we have, setting $G = 2(I+P)^{-1}R$,

$$\|G\|_{L^\infty(\partial\Omega)} \le C_N(\Omega) \quad \text{and} \quad R(z) = \frac{1}{2}\int_{\partial\Omega} G(\sigma)\mu(\sigma,z)\,ds.$$

Comparing the holomorphic and antiholomorphic parts we deduce that, for some complex constant c,

$$R(z) = \frac{1}{4\pi}\int_{\partial\Omega} G(\sigma)\,\frac{\nu}{\sigma-z}\,ds - c \quad \text{and} \quad 0 = \frac{1}{4\pi}\int_{\partial\Omega} G(\sigma)\,\frac{\bar\nu}{\bar\sigma-\bar z}\,ds + c.$$

If $A \in \mathcal{L}(H)$ is an operator with $\overline{W(A)} \subset \Omega$, it is licit to replace z by A and $\bar{z}$ by A^* in the previous relations. After adding we obtain that

$$R(A) = \frac{1}{2} \int_{\partial\Omega} G(\sigma) \otimes \mu(\sigma, A)\, ds,$$

and, using Lemma 2.1, that

$$\|R(A)\| \leq \sup_{z \in \Omega} \|G(z)\|.$$

Therefore we deduce $C_{cb}(\Omega) \leq C_N(\Omega)$.

In order to estimate the constant $C_N(\Omega)$ it is interesting to introduce another constant

$$D_N(\Omega) = \sup\{2 \inf_{c \in \mathbb{C}} \|(I+P)^{-1} r - c\|_{L^\infty(\partial\Omega)}\,; r \text{ rational function, } |r| \leq 1 \text{ in } \Omega\}.$$

It is clear that $D_N(\Omega) \leq C_N(\Omega)$, $P1 = 1$ and $\|Pr\|_{L^\infty(\partial\Omega)} \leq \|r\|_{L^\infty(\partial\Omega)}$. From the relation

$$2(I+P)^{-1} r = r + (I-P)\big((I+P)^{-1} r - c\big), \quad \forall c \in \mathbb{C},$$

we deduce $C_N(\Omega) \leq 1 + D_N(\Omega)$. Therefore we have

$$C_{cb}(\Omega) \leq C_N(\Omega) \leq 1 + D_N(\Omega). \tag{4.3}$$

For an operator $M \in \mathcal{L}(C(\partial\Omega))$ we introduce the norm and semi-norm

$$\begin{aligned} \|M\|_\infty &= \sup\{\|Mf\|_{L^\infty(\partial\Omega)}\,; f \in C(\partial\Omega),\ |f| \leq 1 \text{ in } \Omega\}, \\ \|M\|_{\text{osc}} &= \sup\{\inf_{c \in \mathbb{C}} \|Mf - c\|_{L^\infty(\partial\Omega)}\,; f \in C(\partial\Omega),\ |f| \leq 1 \text{ in } \Omega\}. \end{aligned}$$

Recall that $\operatorname{osc}(f) := 2 \inf_{c \in \mathbb{C}} \|f - c\|_{L^\infty(\partial\Omega)}$ is called the *oscillation* of f on $\partial\Omega$.

We clearly have $C_N(\Omega) \leq 2\,\|(I+P)^{-1}\|_\infty$ and $D_N(\Omega) \leq 2\,\|(I+P)^{-1}\|_{\text{osc}}$.

Remark 4.1. If $\Omega = D$ is a disk of center ω, then simple calculations show that $(Pf)(z) = f(\omega)$ for all harmonic functions f in Ω, with $f \in C(\overline{\Omega})$. We get $(2(I+P)^{-1} f)(z) = 2f(z) - f(\omega)$; therefore

$$C_{cb}(D) \leq C_N(D) = 3 \qquad \text{and} \qquad D_N(D) = 2.$$

The estimate for the first constant is not optimal, since we will see in Theorem 5.1 that $C(D) = C_{cb}(D) = 2$.

Remark 4.2. It is known [25] that, if Ω is not a triangle nor a quadrilateral, then $\|P\|_{\text{osc}} < 1$. This gives the estimate $D_N(\Omega) \leq 2(1 - \|P\|_{\text{osc}})^{-1}$. Furthermore, using the notations

$$\begin{aligned} R_\Omega &:= \sup\{\text{radii of circles which intersect } \partial\Omega \text{ in at least 3 points}\}, \\ L_\Omega &:= \text{perimeter of } \Omega, \end{aligned}$$

we have $\|P\|_{\text{osc}} \leq 1 - \frac{L_\Omega}{2\pi R_\Omega}$ (some smoothness assumptions are mentionned in [25], but we think that they can be avoided). We obtain the estimates

$$D_N(\Omega) \leq \frac{4\pi R_\Omega}{L_\Omega} \qquad \text{and} \qquad C_{cb}(\Omega) \leq 1 + \frac{4\pi R_\Omega}{L_\Omega}.$$

These estimates are not useful if Ω is a polygon since then $R_\infty = +\infty$. For an ellipse (E) with major axis $2a$ and minor axis $2b$, it can be computed that $R_E = a^2/b$; therefore $D_N(E) \leq 2\pi a/b$.

Remark 4.3. The modern proof [17] of the invertibility of $I+P$, which works for any bounded convex domain, follows from the inequality $\|P^2\|_{\rm osc} < 1$. Then, from the relation $(I+P)^{-1} = (I-P)(I-P^2)^{-1}$, we deduce

$$D_N(\Omega) \leq 2\,\|I-P\|_\infty (1-\|P^2\|_{\rm osc})^{-1}.$$

We are not acquainted with a translation of this inequality in simple geometric terms. This was done in another approach developped in [11] which is based on the study of the iterate P^3. Written with the present notations, B. Delyon and F. Delyon have obtained in [11] the estimate

$$C_N(\Omega) \leq 3 + \Big(\frac{2\pi\delta^2}{|\Omega|}\Big)^3, \quad \delta \text{ diameter of } \Omega, \quad |\Omega| \text{ area of } \Omega.$$

4.2 Extension to unbounded domains

The previous developments admit an extension to unbounded domains.

Theorem 4.1. *We assume that the convex domain $\Omega \neq \mathbb{C}$ contains a sector of angle 2α, with $0 < \alpha \leq \pi/2$. Then the C. Neumann problem (4.1) is well posed and we have the estimates*

$$C_N(\Omega) \leq \frac{\pi}{\alpha} \quad \text{and} \quad C_{cb}(\Omega) \leq \frac{\pi-\alpha}{\alpha}.$$

Proof. a) As above, we have

$$\|P\|_\infty \leq \frac{1}{\pi} \sup_{z\in\partial\Omega} \int_{\partial\Omega} d\arg(\sigma - z) = \frac{\pi - 2\alpha}{\pi} < 1.$$

We deduce that $(I+P)$ is invertible from $C(\partial\Omega)$ into itself. Thus the C. Neumann problem has a unique solution and we have

$$\|r\|_{L^\infty(\partial\Omega)} \geq \tfrac{1}{2}(1 - \|P\|_\infty)\,\|g\|_{L^\infty(\partial\Omega)},$$

which implies

$$C_N(\Omega) \leq \frac{2}{1-\|P\|} \leq \frac{\pi}{\alpha}.$$

Remark. For $\varepsilon > 0$ we consider the Banach space $X_\varepsilon := \{f \in C(\partial\Omega)\,;\|f\|_\varepsilon < +\infty\}$, where $\|f\|_\varepsilon := \sup_{z\in\partial\Omega}\big((1+|z|)^\varepsilon|f(z)|\big)$. We set

$$\gamma_\varepsilon := \frac{1}{\pi} \sup_{z\in\partial\Omega} \int_{\partial\Omega_z} \frac{(1+|z|)^\varepsilon}{(1+|\sigma|)^\varepsilon}\, d\arg(\sigma - z).$$

When $\gamma_\varepsilon < +\infty$, using the positivity of P, it is easily seen that P acts from X_ε into itself and that the corresponding induced operator norm is γ_ε. It can be seen that $\limsup_{\varepsilon\to 0}\gamma_\varepsilon \leq 1 - 2\alpha/\pi$. Therefore $I+P$ is invertible in X_ε for ε small enough. In particular, if r is a rational function bounded in Ω and satisfying $r(\infty) = 0$, then the corresponding g in (4.1) belongs to some space X_ε and thus it satisfies an estimate of the form $|g(z)| \leq C(1+|z|)^{-\varepsilon}$.

b) We consider now a rational function R with matrix values and satisfying $\|R(z)\| \leq 1$ in Ω. The corresponding G in the matrix-valued version of (4.1) satisfies $\|G\|_{L^\infty(\partial\Omega)} \leq \pi/\alpha$.

We assume initially that $R(\infty) = 0$. ¿From the previous remark $\|G(\sigma)\| \leq C(1+|\sigma|)^{-\varepsilon}$; this will insure normal convergence in the forthcoming integrals. Then (4.1) implies

$$R(z) = \frac{1}{4\pi}\int_{\partial\Omega} G(\sigma)\,\frac{\nu}{\sigma - z}\,ds, \quad 0 = \frac{1}{4\pi}\int_{\partial\Omega} G(\sigma)\,\frac{\bar\nu}{\bar\sigma - \bar z}\,ds.$$

That justifies to replace z and $\bar{z}$ by A and A^*, if the operator A satisfies $\overline{W}(A) \subset \Omega$. We get

$$R(A) = \frac{1}{2} \int_{\partial\Omega} G(\sigma) \otimes \mu(\sigma, A)\, ds.$$

Note that this equality is still true for any constant R (with the corresponding constant G). Therefore it remains valid without the restriction $R(\infty) = 0$. Then the theorem follows from

$$\|R(A)\| \leq \tfrac{1}{2}\, \|G\|_{L^\infty(\partial\Omega)}\, \|\int_{\partial\Omega} \mu(\sigma, A)\, ds\| = \frac{\pi - \alpha}{\pi}\, \|G\|_{L^\infty(\partial\Omega)} \leq \frac{\pi - \alpha}{\alpha}.$$

□

Remark 4.4.(comparing the bounds) Using Theorem 3.2 and the general bound from [8], we obtain

$$C_{cb}(\Omega) \leq \min\Big(\frac{\pi - \alpha}{\alpha}, 1 + \frac{2}{\pi} \int_{\alpha}^{\pi/2} \frac{\pi - x + \sin x}{\sin x}\, dx, 57\Big).$$

The first bound is the best one if $0.346\,\pi \leq \alpha \leq 0.5\,\pi$ and the second if $4\,10^{-13}\pi \leq \alpha \leq 0.345\,\pi$. The last has to be used only if $0 < \alpha \leq 4\,10^{-13}\pi$.

The result *a fortiori* holds if Ω is a sector of angle 2α, but in this case it is possible to obtain more precise estimates.

Theorem 4.2. *We assume that the convex domain is a sector S_α of angle 2α, with $0 < \alpha \leq \pi/2$. Then we have the following estimates :*

$$C_N(S_\alpha) \leq 2 - \frac{2}{\pi} \log\tan\Big(\frac{\alpha\,\pi}{4(\pi - \alpha)}\Big), \qquad C_N(S_\alpha) = \tfrac{2}{\pi} \log \tfrac{1}{\alpha} + O(1)$$
$$\text{and} \quad C_{cb}(S_\alpha) \leq \frac{\pi - \alpha}{\pi}\, C_N(S_\alpha).$$

Proof. Without loss of generality we can assume that $S_\alpha = \{z \in \mathbb{C}\,; z \neq 0 \text{ and } |\arg z| < \alpha\}$. The proof is based on the following relations

$$g(e^{t+i\alpha}) = 2\, r(e^{t+i\alpha}) - \int_{\mathbb{R}} r(e^{s-i\alpha})\, q(t-s)\, ds,$$
$$g(e^{t-i\alpha}) = 2\, r(e^{t-i\alpha}) - \int_{\mathbb{R}} r(e^{s+i\alpha})\, q(t-s)\, ds,$$
$$\text{where} \quad q(t) = \frac{i}{\pi} \frac{d}{dt} \log \frac{e^{2i\alpha\nu} - e^{-t\nu}}{1 + e^{-t\nu}}, \qquad \nu = \frac{\pi}{2(\pi - \alpha)}.$$

We refer to [10] for all details of these computations. We deduce

$$\|g\|_{L^\infty(S_\alpha)} \leq \big(2 + \|q\|_{L^1(\mathbb{R})}\big)\, \|r\|_{L^\infty(S_\alpha)};$$

therefore

$$C_N(S_\alpha) \leq 2 + \|q\|_{L^1(\mathbb{R})} = 2 - \frac{2}{\pi} \log\tan\Big(\frac{\alpha\,\pi}{4(\pi - \alpha)}\Big).$$

For the function $r(z) = \frac{1-z^{\pi/2\alpha}}{1+z^{\pi/2\alpha}}$ we have $\|r\|_{L^\infty} = 1$. Therefore

$$\begin{aligned} C_N(S_\alpha) \geq \|g\|_{L^\infty} \geq |g(e^{i\alpha})| &= \Big|\int_{\mathbb{R}} r(e^{s-i\alpha})\, q(-s)\, ds\Big| + O(1) \\ &= \Big|\int_{\mathbb{R}} \frac{1+ie^{s\pi/2\alpha}}{1-ie^{s\pi/2\alpha}}\, q(-s)\, ds\Big| + O(1) \\ &= \Big|\int_{\mathbb{R}} \operatorname{sign}(s)\, q(-s)\, ds\Big| + O(1) \\ &= \frac{2}{\pi} \log\frac{1}{\alpha} + O(1). \end{aligned}$$

□

Remark 4.5. We know from [9] or [8] that $C_{cb}(S_\alpha)$ is uniformly bounded, while we have $\lim_{\alpha\to 0} C_N(S_\alpha) = +\infty$. The C. Neumann approach seems to be not appropriate for estimating $C_{cb}(\Omega)$ in the case of flat domains.

4.3 Dilation theorems

In this subsection we assume the invertibility of $I + P$, but the domain Ω may be unbounded. We obtain the following dilation result.

Theorem 4.3. *We assume that the convex domain Ω is such that $I + P$ is an isomorphism of $C(\partial\Omega)$ and that the operator $A \in \mathcal{L}(H)$ satisfies $\overline{W(A)} \subset \Omega$. Then there exists a larger Hilbert space K containing H, and a normal operator N acting on K with spectrum $\sigma(N) \subset \partial\Omega$, such that, for all rational functions r bounded in Ω,*

$$r(A) = P_H g(N)|_H.$$

Here P_H is the orthogonal projection from K onto H and $g = 2(I{+}P)^{-1}r$.

Proof. It follows from (4.1) that

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

The result follows from the Naimark's dilation theorem [22, page 50] which shows the existence of a spectral measure E dilating the regular positive measure $\mu(\sigma, A)\, ds$. □

Remark 4.6. If $\Omega = D$ is the unit disk and $r(z) = z^n$, with $n \geq 1$, we have (see Remark 4.1) $g = 2(I{+}P)^{-1}r = 2\,r$. In this case, the above theorem reduces to the 2-dilation theorem of Berger [5] which states that every $A \in \mathcal{L}(H)$ with $W(A) \subset \overline{D}$ satisfies

$$A^n = 2\, P_H U^n|_H, \qquad \forall\, n \geq 1,$$

for a suitable unitary operator U acting on K.

Corollary 4.4 (T. Kato). *Let f be a rational function such that $f(\infty) = \infty$ and the set $\Omega := \{z \in \mathbb{C} : |f(z)| < 1\}$ is convex. Let $A \in \mathcal{L}(H)$ be a linear operator such that its numerical range satisfies $W(A) \subset \overline{\Omega}$. Then we have $W(f(A)) \subset \overline{D}$.*

Proof. We set $\rho(\bar z) = \overline{1/f(z)}$. Notice that ρ is antiholomorphic out of Ω and $\rho(\bar\sigma) = f(\sigma)$ on $\partial\Omega$. We deduce from the Cauchy formula that, for all $k \geq 1$, and all $z \in \Omega$,

$$\begin{aligned}(f(z))^k &= \int_{\partial\Omega} f^k(\sigma)\,\mu(\sigma,z)\,ds + \frac{1}{2\pi i}\int_{\partial\Omega} f^k(\sigma)\frac{d\bar\sigma}{\bar\sigma - \bar z}\\ &= \int_{\partial\Omega} f^k(\sigma)\,\mu(\sigma,z)\,ds + \frac{1}{2\pi i}\int_{\partial\Omega} \rho^k(\bar\sigma)\frac{d\bar\sigma}{\bar\sigma - \bar z} = \int_{\partial\Omega} f^k(\sigma)\,\mu(\sigma,z)\,ds.\end{aligned}$$

Therefore $g = 2\,f^k$ is the solution of the Neumann problem (4.1) for the data $r = f^k$, i.e. $f^k = (I+P)^{-1} f^k$ (or equivalently $Pf^k = 0$). From the previous theorem we deduce that

$$(f(A))^k = 2\,P_H\,U^k|_H, \qquad \forall\,k \geq 1,$$

where $U = f(N)$ is a normal operator with spectrum $\sigma(U) \subset f(\partial\Omega) \subset \partial D$. Therefore U is a unitary operator and U is a unitary 2-dilation of $f(A)$. Then it follows from [5] that $W(f(A)) \subset \overline{D}$. ☐

Remark 4.7. A domain Ω as in Corollary 4.4 is called a *convex lemniscate*. The spectral mapping theorem stated in the above Corollary, which is due to T. Kato [16], was proved here using a different method. The key point in our proof is that $Pf^k = 0$ for all $k \geq 1$ (compare with [12, Theorem 2.1]).

Remark 4.8. After a first version of this paper has been completed, the recent article [24] was brought to the authors' attention by John McCarthy and Mihai Putinar. There is some overlapping between Subsection 4.3 and some of the results of [24].

5 The similarity approach

5.1 The case of the disk

We first look to the case where Ω is the unit disk $D = \{z \in \mathbb{C}\,;\,|z| < 1\}$.

Theorem 5.1. *In the disk case we have $C_{cb}(D) = C(D,2) = 2$.*

Proof. Let A be an operator with $\overline{W(A)} \subset D$. From the Berger thorem (see Remark 4.6) we know that A admits a 2-unitary dilation. Then a result due to Okubo and Ando [20] states the following :

$$\text{there exists an invertible operator } S \text{ such that } \|S\|\,\|S^{-1}\| \leq 2 \text{ and } \|S\,A\,S^{-1}\| \leq 1. \tag{5.1}$$

Using the von Neumann inequality for the contraction $S\,A\,S^{-1}$ we deduce that

$$\|R(S\,A\,S^{-1})\| \leq 1, \quad \text{for all matrix-valued rational functions with } \|R(z)\| \leq 1 \text{ in } D.$$

We deduce $C_{cb}(D) \leq 2$ from the inequality $\|R(A)\| \leq \|S^{-1}\|\|R(S\,A\,S^{-1})\|\|S\|$.

With the choice $r(z) = z$ and $A = \begin{pmatrix} 0 & 2\\ 0 & 0\end{pmatrix}$ we see that the bound 2 is attained. ☐

5.2 The case of 2×2 matrices

We turn now to the case of 2×2 matrices. In this case it is known that the numerical range is an ellipse whose foci are the eigenvalues. This ellipse is degenerate only if the matrix is normal.

Theorem 5.2. *Let A be a 2×2 (non normal) matrix. Then there exists a conformal function a mapping the ellipse $W(A)$ onto the unit disk D and an invertible matrix S such that $\|S\|\,\|S^{-1}\| \leq 2$ and $\|S\,a(A)\,S^{-1}\| \leq 1$.*

Proof. If the eigenvalues of A are equal, then $W(A)$ is a disk and the result follows from (5.1). Suppose now that A has distinct eigenvalues. Since any matrix is unitary similar to an upper triangular matrix, we can assume that A is upper triangular. Furthermore, it is clear that if the theorem holds for a matrix A, then it also holds for $\lambda A + \beta I$ for any $\lambda \neq 0$ and $\beta \in \mathbb{C}$. Finally, we only have to look at matrices of the form $A = \begin{pmatrix} 1 & \gamma \\ 0 & -1 \end{pmatrix}$ and we can furthermore assume that $\gamma > 0$. Then $W(A)$ is the ellipse of foci 1, -1 and minor axis γ (see [13] for instance). If $\rho > 1$ is chosen such that $\gamma = \rho - 1/\rho$, then the major axis is $\rho + 1/\rho$. Also (see [15]), the function

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

where t_n denotes the n^{th} Chebyshev polynomial, is the Riemann conformal function mapping (the interior of) $W(A)$ onto D. Notice that $a(1) = -a(-1)$; thus $a(A) = a(1)A$.

We choose now

$$S = \begin{pmatrix} 1 + a(1)^2 & a(1)^2\rho - 1/\rho \\ 0 & a(1)(\rho + 1/\rho) \end{pmatrix};$$

then we have

$$B := Sa(A)S^{-1} = \begin{pmatrix} a(1) & 1-a(1)^2 \\ 0 & -a(1) \end{pmatrix}.$$

It is easy to verify that $\|B\| = 1$. Some simple computations show that the quantity $\|S\|\,\|S^{-1}\|$ is the largest root of the equation

$$X^2 - \frac{1+\rho^2 a(1)^2}{\rho\, a(1)}\,X + 1 = 0,$$

which is $X = \rho\, a(1) = 2\,\exp\big(-\sum_{n\geq 1}\frac{(-1)^{n+1}}{n}\,\frac{2}{1+\rho^{4n}}\big)$. It is easy to verify that $1 < X < 2$. □

Corollary 5.3. *In the two dimensional case we have $\sup_\Omega C_{cb}(\Omega, 2) = \sup_\Omega C(\Omega, 2) = 2$.*

Proof. We know from the previous subsection that $C(D, 2) = 2$, thus it is sufficient to prove, for any 2×2 matrix A, that $C_{cb}(W(A), 2) \leq 2$. Let R be a rational function bounded in $W(A)$; we set $Q(z) = R(a^{-1}(z))$ and $B = S\,a(A)\,S^{-1}$. We have

$$R(A) = Q(a(A)) = S^{-1}Q(B)S.$$

Using the von Neumann inequality for the contraction B, we obtain

$$\|R(A)\| \leq 2\,\|Q(B)\| \leq 2\,\sup_{z\in D}\|Q(z)\| = 2\sup_{\zeta\in W(A)}\|R(\zeta)\|;$$

thus $C_{cb}(W(A), 2) \leq 2$. □

Remark 5.1. The equality $C(\Omega, 2) = C_{cb}(\Omega, 2)$ follows also from [23].

5.3 The case of ellipses

Theorem 5.4. *Let $\mathcal{E}$ be (the interior of) an ellipse in $\mathbb{C}$ with foci μ_1 and μ_2 and let γ be the length of the minor axis. Let A be an operator in $\mathcal{L}(H)$. Then $W(A) \subset \overline{\mathcal{E}}$ if and only if there exits an isometry V from H into $H \oplus H$ such that*

$$A = V^*(E \otimes I)V, \qquad \text{where } E = \begin{pmatrix} \mu_1 & \gamma \\ 0 & \mu_2 \end{pmatrix}.$$

Proof. The *if* part is easily verified. For the *only if* part we can assume without loss of generality that $\mathcal{E} = \{x+iy\,; \cos^2\theta\, x^2 + y^2 < 1\}$ for some real θ, i.e. $\mu_1 = -\mu_2 = \tan\theta$ and $\gamma = 2$. Then we write

$$M = \tfrac{1}{2}(A + A^*), \quad N = \tfrac{1}{2i}(A - A^*); \quad \text{thus} \quad A = M + iN.$$

It is easily seen that

$$W(A) \subset \overline{\mathcal{E}} \quad \Longleftrightarrow \quad W(\cos\theta M + iN) \subset \overline{D}.$$

According to a result of Ando [1], there exist a unitary operator U and a self-adjoint operator B such that

$$\cos\theta M + iN \;=\; 2\,\sin B\; U\; \cos B.$$

Therefore

$$\begin{aligned} \cos\theta M &= \sin B\; U\; \cos B + \cos B\; U^*\; \sin B, \\ iN &= \sin B\; U\; \cos B - \cos B\; U^*\; \sin B. \end{aligned}$$

We deduce

$$M + iN = \frac{1}{\cos\theta} W^* \begin{pmatrix} 0 & 1+\cos\theta \\ 1-\cos\theta & 0 \end{pmatrix} W, \quad \text{with} \quad W = \begin{pmatrix} \sin B \\ U\cos B \end{pmatrix}.$$

It is clear that W is an isometry. Now we remark that

$$\begin{pmatrix} 0 & 1+\cos\theta \\ 1-\cos\theta & 0 \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \begin{pmatrix} \sin\theta & 2\cos\theta \\ 0 & -\sin\theta \end{pmatrix} \begin{pmatrix} \cos\frac{\theta}{2} & \sin\frac{\theta}{2} \\ -\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}.$$

The theorem follows by taking

$$V = \begin{pmatrix} \cos\frac{\theta}{2} & \sin\frac{\theta}{2} \\ -\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} W.$$

□

Remark 5.2. Theorem 5.4 is useful for estimates involving polynomials of degree one with matrix coefficients. Noticing that for such a polynomial P we have

$$P(A) = V^*(P(E) \otimes I_H)V,$$

we deduce

$$\|P(A)\| \le 2 \sup_{z\in\mathcal{E}} \|P(z)\|,$$

for all polynomials P of degree one with matrix coefficients.

References

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

[2] W. Arendt, *Semigroups and Evolution Equations : Functional Calculus, Regularity and Kernel Estimates*, to appear in Handbook of Differential Equations : Evolutionary Differential Equations. Available from `http://www.mathematik.uni-ulm.de/m5/arendt/`

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

[4] B. Beckerman, S.A. Goreinov, E.E. Tyrtyshnikov, Some remarks on the Elman estimate for GMRES, Preprint.

[5] C. A. Berger, J. G. Stampfli, Mapping theorems for the numerical range, *Amer. J. Math.* **89**, 1047–1055, 1967.

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

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

[8] M. Crouzeix, Numerical range and hilbertian functional calculus, Preprint. Available from `http://perso.univ-rennes1.fr/michel.crouzeix/`

[9] M. Crouzeix, B. Delyon, Some estimates for analytic functions of strip or sectorial operators, *Arch. Math.*, **81**, 553-566, 2003.

[10] M. Crouzeix, B. Delyon, Sur le problème de C. Neumann dans un secteur, Preprint. Available from `http://perso.univ-rennes1.fr/michel.crouzeix/`

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

[12] P. Ebenfelt, D. Khavinson, H. S. Shapiro, An inverse problem for the double layer potential, *Comput. Methods Funct. Theory* **1**, no. 2, 387–401, 2001.

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

[14] M. Haase, *The Functional Calculus for Sectorial Operators*, book manuscript, 2004. Available from `http:// www.mathematik.uni-ulm.de/m5/haase`

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

[16] T. Kato, Mapping theorems for the numerical range, *Proc. Japan Acad.*, vol. 41, no 8, 652–655, 1965.

[17] J. Král, *Integral operators in potential theory*, Lecture Notes in Mathematics 823, Springer, 1980.

[18] C. Le Merdy, Similarities of ω-accretive operators, CMA Proceedings (Canberra), 2002.

[19] C. Neumann, *Über die Methode des arithmetischen Mittels*, Hirzel, Leipzig, 1887 (erste Abhandlung), 1888 (zweite Abhandlung).

[20] K. Okubo, T. Ando, Constants related to operators of class C_ρ, *Manuscripta Math.* **16**, no 4, 385-394, 1975.

[21] V. Paulsen, Every completely polynomially bounded operator is similar to a contraction, *J. Funct. Anal.*, vol. **55** , no. 1, 1–17, 1984.

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

[23] V. Paulsen, K-spectral values for finite matrices, *J. Operator Theory*, vol. **18**, no 2, 249–263, 1987.

[24] M. Putinar, S. Sandberg, A skew normal dilation on the numerical range of an operator, to appear in *Math. Ann.* Available from http://www.math.ucsb.edu/%7Emgscharl/preprints/2002preprints.htm

[25] G. Schober, Neumann's lemma, *Proc. Amer. Math. Soc.* vol. **19**, 306–311, 1968.

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

C. Badea
Département de Mathématiques, UMR CNRS no. 8524,
Université de Lille I, F–59655 Villeneuve d'Ascq, France
catalin.badea@univ-lille1.fr

M. Crouzeix and B. Delyon
Institut de Recherche Mathématique de Rennes, UMR CNRS no. 6625,
Université de Rennes 1, Campus de Beaulieu, 35042 RENNES Cedex, France
michel.crouzeix@univ-rennes1.fr
bernard.delyon@univ-rennes1.fr