

Image numérique et calcul fonctionnel hilbertien

Problèmes ouverts

Michel CROUZEIX

janvier 2006

Abstract

Ce document n'est pas destiné à devenir un article. C'est plutôt un document de travail écrit pour présenter assez rapidement les résultats existants sur le sujet, et les nombreux problèmes et sous-problèmes qui restent ouverts. C'est aussi un bric à brac où j'ai voulu inclure quelques pistes que j'ai explorées, et des réflexions diverses, en espérant qu'elles puissent être utiles. A priori c'est un document destiné à être évolutif. En dehors de moi, relativement peu de personnes¹ ont réellement abordé ce domaine, je pense qu'il reste encore beaucoup de résultats à obtenir et sans doute d'applications à trouver. N'hésitez pas à me signaler mes erreurs, me faire connaître vos commentaires ou vos progrès sur ce sujet.

1 Introduction.

Soit $A \in \mathbb{C}^{d,d}$ une matrice carrée à coefficients complexes. On lui associe le sous-ensemble du plan complexe $W(A) := \{\langle Av, v \rangle; \|v\| = 1, v \in \mathbb{C}^d\}$; $\langle \cdot, \cdot \rangle$ désigne le produit scalaire usuel sur \mathbb{C}^d , $\|\cdot\|$ la norme associée. Cet ensemble $W(A)$ est appelé l'image numérique de A , c'est un ensemble convexe (théorème de Toeplitz et Hausdorff). Le point central de ce document est le résultat suivant [4], dont l'énoncé est très élémentaire

Il existe une meilleure constante \mathcal{Q} telle que

$$\|p(A)\| \leq \mathcal{Q} \sup_{z \in W(A)} |p(z)|, \quad (1)$$

- pour tout polynôme $p : \mathbb{C} \rightarrow \mathbb{C}$,
- pour toute matrice $A \in \mathbb{C}^{d,d}$, pour toute valeur de d .

De plus, on a l'estimation $2 \leq \mathcal{Q} \leq 11.08$.

Précisons que, pour une matrice $M \in \mathbb{C}^{d,d}$, on utilise la norme $\|M\| := \sup\{\|Mv\|; v \in \mathbb{C}^d, \|v\| = 1\}$, norme d'opérateur induite par la structure hilbertienne de \mathbb{C}^d . Rappelons que, si la matrice A était normale, on aurait l'estimation bien meilleure $\|p(A)\| \leq \sup_{z \in \sigma(A)} |p(z)|$, où $\sigma(A)$ désigne le spectre de A : on vérifie aisément que $\sigma(A) \subset W(A)$, d'où a fortiori la majoration (1) dans ce cas particulier. L'intérêt de (1) est de permettre de sortir du cadre restreint des matrices normales.

Il est remarquable que le résultat ci-dessus soit encore valide en version tensorielle (on dit aussi en version complètement bornée). Plus précisément on a

Il existe une meilleure constante \mathcal{Q}_{cb} telle que

$$\|P(A)\| \leq \mathcal{Q}_{cb} \sup_{z \in W(A)} \|P(z)\|, \quad (2)$$

¹B. Delyon, C. Badea et B. Beckermann pour des résultats, Chr. Le Merdy et M. Haase pour des applications

- pour tout polynôme $P : \mathbb{C} \rightarrow \mathbb{C}^{m,n}$, pour toute valeur de m et n ,
 - pour toute matrice $A \in \mathbb{C}^{d,d}$, pour toute valeur de d .
- De plus, on a l'estimation $2 \leq \mathcal{Q}_{cb} \leq 11.08$.

Ici le polynôme P est à valeurs matricielles $P(z) = (p_{ij}(z))$; la matrice $P(A) \in \mathbb{C}^{md,nd}$ est la matrice formée de $m \times n$ blocs $d \times d$, le bloc i, j étant la matrice $p_{ij}(A)$.

Le côté surprenant de ces majorations est que les constantes \mathcal{Q} et \mathcal{Q}_{cb} existent indépendamment à la fois, de la matrice A , de sa dimension, du degré des polynômes utilisés, et de plus de m et n pour \mathcal{Q}_{cb} . Ces caractères d'uniformité permettent d'étendre ces estimations, d'une part à n'importe quel opérateur linéaire continu $A \in \mathcal{L}(H)$ sur un espace de Hilbert H (voire à des opérateurs non bornés), d'autre part à toute fonction p (resp. P) continue sur $\overline{W(A)}$ et holomorphe en son intérieur. Nous renvoyons à [4] pour ces extensions ainsi que pour l'exposé de quelques applications.

En contraste avec l'énoncé très élémentaire de (1), la démonstration que j'en ai obtenue est pour le moins compliquée, assez technique, et peu naturelle. Elle aboutit à une surestimation, sans doute très importante, de la valeur de la constante : $\mathcal{Q} \leq 11.08$; elle a cependant l'avantage de rester valide sous forme tensorielle. On a donc

$$2 \leq \mathcal{Q} \leq \mathcal{Q}_{cb} \leq 11.08.$$

Quelques problèmes ouverts. Ma conjecture est que $\mathcal{Q} = \mathcal{Q}_{cb} = 2$. J'expliquerai plus loin quelques raisons qui me font croire à cette éventualité. Deux sous-problèmes qui en découlent sont *Est-ce que $\mathcal{Q} = 2$? Est-ce que $\mathcal{Q} = \mathcal{Q}_{cb}$?* Si cette dernière relation s'avérait vraie, il serait intéressant de comprendre pourquoi cela se passe ainsi dans ce cas, alors que la conjecture de Halmos, *borné implique complètement borné*, s'est avérée fausse.

Je crains que les problèmes que je viens de mentionner ne soient très difficiles. Pour obtenir des problèmes peut-être plus accessibles, on peut définir

$$\mathcal{Q}(d) := \sup_{A,p} \{ \|p(A)\| ; A \in \mathbb{C}^{d,d}, p : \mathbb{C} \rightarrow \mathbb{C} \text{ polynôme, } |p(z)| \leq 1 \text{ dans } W(A) \},$$

$$\mathcal{Q}_{cb}(d) := \sup_{A,P,m,n} \{ \|P(A)\| ; A \in \mathbb{C}^{d,d}, P : \mathbb{C} \rightarrow \mathbb{C}^{m,n} \text{ polynôme, } \|P(z)\| \leq 1 \text{ dans } W(A) \}.$$

On vérifie aisément que ces constantes sont des fonctions croissantes de la dimension d ; de plus on a $\mathcal{Q} = \sup_d \mathcal{Q}(d)$ et $\mathcal{Q}_{cb} = \sup_d \mathcal{Q}_{cb}(d)$. J'ai réussi à montrer [1] que $\mathcal{Q}(2) = \mathcal{Q}_{cb}(2) = 2$, mais j'ai échoué pour les relations $\mathcal{Q}(3) = \mathcal{Q}_{cb}(3)$ et $\mathcal{Q}(3) = 2$; j'ai testé cette dernière égalité par des essais numériques assez délicats, et je n'ai pas réussi à la contredire. Bien évidemment les mêmes questions sont ouvertes pour les valeurs de $d > 3$.

Plus généralement trouver une démonstration de nature différente de celle que j'ai proposée pour les estimations (1) et (2) pourrait être très utile.

Obtient-on des constantes plus petites si on se limite aux matrices A à coefficients réels ?

Peut-on trouver un résultat analogue pour deux matrices qui commutent ?

2 Problèmes associés à une matrice

Soit $A \in \mathbb{C}^{d,d}$ une matrice carrée, et Ω un ouvert borné convexe. On définit

$$\begin{aligned}\psi_\Omega(A) &:= \sup_p \{ \|p(A)\| ; p : \mathbb{C} \rightarrow \mathbb{C} \text{ polynôme, } |p(z)| \leq 1 \text{ dans } \Omega \}, \\ \psi_{cb,\Omega}(A) &:= \sup_{P,m,n} \{ \|P(A)\| ; P : \mathbb{C} \rightarrow \mathbb{C}^{m,n} \text{ polynôme, } \|P(z)\| \leq 1 \text{ dans } \Omega \}, \\ \psi(A) &:= \sup_p \{ \|p(A)\| ; p : \mathbb{C} \rightarrow \mathbb{C} \text{ polynôme, } |p(z)| \leq 1 \text{ dans } W(A) \}, \\ \psi_{cb}(A) &:= \sup_{P,m,n} \{ \|P(A)\| ; P : \mathbb{C} \rightarrow \mathbb{C}^{m,n} \text{ polynôme, } \|P(z)\| \leq 1 \text{ dans } W(A) \}.\end{aligned}$$

Il est clair que $\psi_\Omega(A)$ et $\psi_{cb,\Omega}(A)$ sont des fonctions décroissantes de Ω (pour la relation d'ordre inclusion), de plus

$$\psi(A) = \sup_\Omega \{ \psi_\Omega(A) ; \Omega \supset W(A) \}, \quad \psi_{cb}(A) = \sup_\Omega \{ \psi_{cb,\Omega}(A) ; \Omega \supset W(A) \}.$$

Notons de plus que, d'après [13], cela ne modifie pas les valeurs de $\psi_{cb,\Omega}(A)$ ni de $\psi_{cb}(A)$ si on se limite dans leurs définitions à prendre $m = n = d$.

On a

$$\mathcal{Q}(d) = \max\{ \psi(A) ; A \in \mathbb{C}^{d,d} \}, \quad \mathcal{Q}_{cb}(d) = \max\{ \psi_{cb}(A) ; A \in \mathbb{C}^{d,d} \}.$$

Ces relations sont évidentes si on remplace *max* par *sup*. On montre assez facilement que les bornes sont effectivement atteintes.

Remarque. Les définitions précédentes restent valables si A est un opérateur linéaire borné sur un espace de Hilbert. Rappelons que la conjecture de Halmos était $\psi_D(A) < +\infty$ implique $\psi_{cb,D}(A) < +\infty$ où D est le disque unité. Gilles Pisier a montré [11] qu'elle était fautive (bien sûr en dimension infinie). Il en résulte que, même en dimension finie, la relation $\psi_\Omega(A) = \psi_{cb,\Omega}(A)$ ne peut pas être vraie de manière générale. Cependant le problème $\psi(A) = \psi_{cb}(A)$ est, pour moi, ouvert.

Pour l'étude de ces quantités, on peut ramener l'étude sur Ω à celle sur le disque unité, en faisant intervenir une transformation conforme a de Ω sur D . On a alors (cf. [3] par exemple) $\psi_\Omega(A) = \psi_D(a(A))$ et $\psi_{cb,\Omega}(A) = \psi_{cb,D}(a(A))$; de plus, si B est une matrice $d \times d$ ayant ses valeurs propres intérieures au disque unité, $\psi_D(B)$ est atteint pour un produit de Blaschke à au plus $d - 1$ termes. Plus précisément on a

$$\psi_D(B) := \sup_{\zeta_j} \{ \|g(B)\| ; g(z) = \prod_{j=1}^r \frac{z - \zeta_j}{1 - \bar{\zeta}_j z}, \zeta_1, \dots, \zeta_r \in D, r \leq d - 1 \}.$$

Pour la quantité complètement bornée associée, on peut se ramener aussi à un problème d'optimisation en utilisant la caractérisation due à Paulsen [9]

$$\psi_{cb,D}(B) := \min_S \{ \|S\| \|S^{-1}\| ; S \in \mathbb{C}^{d,d}, \|S^{-1}BS\| \leq 1 \}.$$

Quelques problèmes ouverts. Trouver une méthode numérique effective pour calculer $\psi_D(B)$. Pour $d = 2$ on peut calculer la solution explicitement, mais pour $d \geq 3$ on ne sait pas si la méthode numérique ne fournit pas la valeur d'un maximum local au lieu du vrai maximum. De même je ne connais pas de méthode efficace pour trouver une matrice S réalisant le minimum dans la caractérisation de $\psi_{cb,D}(B)$, la démonstration de son existence étant pour le moins non

constructive : utilisation répétée du théorème de Hann-Banach, ainsi que plusieurs passages au quotient...

Est-ce-que la fonction $A \mapsto \psi(A)$ est continue lorsque la matrice (ou plus généralement l'opérateur) n'est pas de la forme $A = \lambda I$? On voit facilement que c'est le cas lorsque le spectre de A ne possède pas de point appartenant à la frontière de $W(A)$, par utilisation de la formule de Cauchy sur cette frontière. Il est clair que la fonction est semi-continue inférieurement. Elle n'est pas continue en λI puisque

$$\psi \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = 1 \text{ et } \psi \begin{pmatrix} \lambda & \varepsilon \\ 0 & \lambda \end{pmatrix} = 2.$$

Même question pour $\psi_{cb}(\cdot)$

On définit maintenant (pour p holomorphe dans D)

$$\psi_{d,k}(B) := \sup_p \{ \|p(B)\| ; p : \mathbb{C} \rightarrow \mathbb{C}, |p(z)| \leq 1 \text{ dans } \Omega, \text{ ayant au plus } k \text{ zéros dans } D \}.$$

Est-ce-que $\psi_{d,k}(B)$ (avec $k \geq 1$) est atteint pour p produit de Blaschke à, au plus, k termes ?

3 Constantes associées à un domaine convexe

Soit $\Omega \neq \mathbb{C}$ un ouvert convexe, non vide, borné ou non, du plan complexe. On définit les constantes

$$C(\Omega, d) := \sup_{A,r} \{ \|r(A)\| ; A \in \mathbb{C}^{d,d}, W(A) \subset \Omega, r : \mathbb{C} \rightarrow \mathbb{C}, |r(z)| \leq 1, \forall z \in \Omega \},$$

$$C_{cb}(\Omega, d) := \sup_{A,R,m,n} \{ \|R(A)\| ; A \in \mathbb{C}^{d,d}, W(A) \subset \Omega, R : \mathbb{C} \rightarrow \mathbb{C}^{m,n}, \|R(z)\|_2 \leq 1, \forall z \in \Omega \}.$$

$$C(\Omega) := \sup_d C(\Omega, d), \quad C_{cb}(\Omega) := \sup_d C_{cb}(\Omega, d).$$

Dans ces définitions, r et R désignent des fractions rationnelles. (Nous avons fait ce choix pour traiter le cas intéressant des domaines non bornés, mais dans le cas où Ω est borné on pourrait se restreindre à choisir r et R polynomiaux, sans changer les bornes supérieures. De même on pourrait remplacer la condition $W(A) \subset \Omega$ par $W(A) \subset \overline{\Omega}$).

On a bien évidemment

$$\mathcal{Q}(d) = \sup_{\Omega} C(\Omega, d), \quad \mathcal{Q} = \sup_{\Omega} C(\Omega), \quad \mathcal{Q}_{cb}(d) = \sup_{\Omega} C_{cb}(\Omega, d), \quad \mathcal{Q}_{cb} = \sup_{\Omega} C_{cb}(\Omega)$$

Remarques.

1) Les constantes définies précédemment ne dépendent que de d et de la forme de Ω . Plus précisément, si φ est une similitude : $\varphi(z) = a + bz$, ou une anti-similitude : $\varphi(z) = a + b\bar{z}$, $a, b \in \mathbb{C}$, $b \neq 0$, on a $C(\Omega) = C(\varphi(\Omega)), \dots, C_{cb}(\Omega, d) = C_{cb}(\varphi(\Omega), d)$.

2) D'après un résultat classique dû à J. von Neumann [12], on a $C(\Omega) = 1$ lorsque Ω est un demi-plan ; dès que la notion de complètement bornée est apparue, on a remarqué que l'on a aussi dans ce cas $C_{cb}(\Omega) = 1$.

3) On a clairement $C(\Omega) \leq C_{cb}(\Omega)$ et $C(\Omega, d) \leq C_{cb}(\Omega, d)$. De plus, ces deux dernières constantes sont des fonctions croissantes de d .

En dehors du cas bien connu du demi-plan, le premier résultat est assez récent. Dans le joli article [7] il est montré que l'on a $C(\Omega) < +\infty$, pour tout ouvert convexe borné Ω ; dans [1] nous avons montré de plus que cela reste vrai en version complètement bornée, et donné l'estimation

$$C_{cb}(\Omega) \leq 2 + \pi + \inf_{\omega \in \partial\Omega} \text{TV}(\log |\sigma - \omega|);$$

où $\text{TV}(\log |\sigma - \omega|)$ désigne la variation totale de $\log(|\sigma - \omega|)$ lorsque σ parcourt $\partial\Omega$.

Une approche similaire donne aussi

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

lorsque Ω est contenu dans un secteur d'angle 2α , $0 < \alpha \leq \frac{\pi}{2}$. Pour le secteur lui-même $\Omega = S_{\alpha}$ (d'angle $2\alpha \leq \pi$) nous avons une estimation plus précise [5], [1],

$$\frac{\pi \sin \alpha}{2\alpha} \leq C(S_{\alpha}) \leq C_{cb}(S_{\alpha}) \leq \min\left(\frac{\pi - \alpha}{\alpha}, 2 + \frac{2}{\sqrt{3}}, \frac{\pi - \alpha}{\pi} \left(2 - \frac{2}{\pi} \log \text{tg} \left(\frac{\alpha \pi}{4(\pi - \alpha)}\right)\right)\right).$$

Cette borne reste valide dans le cas de la bande $\Omega = S_0$.

Une amélioration récente est due à Bernhard Beckermann (cf. [2]). Il a obtenu l'estimation suivante,

$$C(S_{\alpha}) \leq C_{cb}(S_{\alpha}) \leq 2 - \frac{2\alpha}{\pi} + \frac{2 \cos \alpha}{\pi \sqrt{1 + 2 \cos 2\alpha}} \arccos \left(\frac{\cos(\pi - 2\alpha)}{\cos \alpha} \right), \quad \text{pour } \alpha \in [0, \pi/3].$$

Cette borne est meilleure que la précédente lorsque $\alpha \leq .22\pi$, et reste aussi valable lorsque la frontière de Ω est une branche d'hyperbole d'angle 2α . Dans l'article [6], on montre que $C_{cb}(\mathcal{P}) \leq 4.75$ si \mathcal{P} est une parabole, estimation que l'on vient d'améliorer en $C_{cb}(\mathcal{P}) \leq 2 + 2/\sqrt{3}$. Dans [2] on donne la borne $C_{cb}(\mathcal{E}) \leq 2 + 2/\sqrt{4 - e^2}$ pour une ellipse d'excentricité e .

En dehors du cas du demi-plan, la seule valeur exacte connue est celle, obtenue par Catalin Badea, du disque $C(D) = C_{cb}(D) = 2$, voir [1]. Les autres bornes, et surtout la borne uniforme $C_{cb}(\Omega) \leq 11.08$, sont très pessimistes.

Nous sommes cependant très loin des valeurs exactes. Je suis convaincu de la véracité de

$$\text{Ma conjecture :} \quad \mathcal{Q} = \mathcal{Q}_{cb} = \sup_{\Omega} C_{cb}(\Omega) = 2.$$

Pour le moment j'ai seulement réussi à montrer que [1], $\mathcal{Q}(2) = \mathcal{Q}_{cb}(2) = 2$.

Problèmes ouverts.

- Est ce que l'on a $C(\Omega) = C_{cb}(\Omega)$ pour tout domaine convexe Ω ?
(d'après [10], on sait que $C(\Omega, 2) = C_{cb}(\Omega, 2)$).
- Est ce que $C_{cb}(S_0) \leq 2$?
- Peut-on trouver un résultat d'interpolation, du genre Marcel Riesz, pour majorer $C_{cb}(S_{\alpha})$ à partir de $C_{cb}(S_0)$ et $C_{cb}(S_{\pi/2}) = 1$?
- Est ce que $C(\Omega) = 2$ implique Ω est un disque ? (cela est vrai en dimension 2, voir [3]).
- Est ce que $C(\Omega)$ (resp. $C(\Omega, d)$) est une fonction continue de Ω (par rapport à la distance de Hausdorff) ? (On montre assez facilement qu'elle est semi-continue inférieurement).
Tout au moins, est ce que $C(\Omega) \rightarrow 2$ lorsque Ω tend vers le disque unité ?

- Est ce que $\mathcal{Q} = \sup_{\Omega} C(\Omega)$ (resp. $\mathcal{Q}_{cb} = \sup_{\Omega} C_{cb}(\Omega)$) est atteint pour un Ω ? (C'est en fait le cas pour $\mathcal{Q}(d)$ et $\mathcal{Q}_{cb}(d)$). Est-il atteint pour un Ω symétrique par rapport à l'axe réel ?
- Est ce que $C(\Omega, d)$ (resp. $\sup_{\Omega} C_{cb}(\Omega, d)$) est atteint pour une matrice A ? (Il le serait alors pour une fonction r). Remarquons que c'est le cas pour $C(\Omega, d)$, lorsque Ω est borné et sa frontière analytique [3], la fonction r est alors un produit de Blaschke.
- Lorsque la frontière de Ω est une branche d'hyperbole d'angle 2α , a-t-on l'égalité $C(\Omega, d) = C(S_{\alpha}, d)$?
- Lorsque Ω est symétrique par rapport à l'axe réel, est-ce-que $C(\Omega, d)$ (resp. $C_{cb}(\Omega, d)$) est atteint pour une matrice A à coefficients réels ? Plus généralement quelles relations y-a-t'il entre les propriétés de symétrie de Ω et celles des matrices A qui réalisent $C(\Omega, d)$?
- Trouver une méthode numérique pour calculer la valeur de $C(\Omega, d)$, pour $d = 2, 3, \dots$ et Ω donné. (J'ai seulement réussi à le faire dans le cas de la bande S_0 pour $d \leq 8$; les résultats montrent que $C(S_0, d)$ croît strictement avec d , la limite semble être ≤ 2)...

4 Quelques réflexions sur l'image numérique

Pour une description assez complète de l'image numérique, son historique et ses différentes propriétés, je renvoie à [8]. Nous nous limitons ici à quelques remarques.

L'image numérique est un sous-ensemble convexe du plan complexe. En dehors du cas des matrices 2×2 , et même en petite dimension 3 ou 4, ce n'est pas un objet simple à décrire; mais on peut en obtenir une bonne approximation en remarquant qu'un ensemble convexe est l'intersection des demi-plans tangents qui le contiennent. Plus précisément, pour une matrice $A = B + iC \in \mathbb{C}^{d,d}$, B et C étant auto-adjointes, on pose $P_A(u, v, w) := \det(uB + vC + wI)$, et on note $w_m(u, v)$ la plus grande racine de $P_A(u, v, \cdot) = 0$ (ces racines sont toutes réelles puisque B et C sont auto-adjointes). On peut alors montrer que

$$W(A) = \{z = x + iy; x \cos \alpha + y \sin \alpha + w_m(\cos \alpha, \sin \alpha) \leq 0, \text{ pour tout } \alpha \in [0, 2\pi]\}.$$

On obtient ainsi une approximation (par l'extérieur) de $W(A)$ en calculant un nombre fini de valeurs de $w_m(\cdot, \cdot)$.

Notons que l'approche par la voie tangentielle est beaucoup plus simple que l'approche cartésienne. D'après ce qui précède la frontière de $W(A)$ est une partie de la courbe algébrique d'équation tangentielle $P_A(u, v, w) = 0$. Le polynôme P_A étant de degré d , la courbe est dite de classe d . On peut déduire classiquement de l'équation tangentielle, l'équation cartésienne de cette courbe, mais il apparaît que celle-ci est génériquement de degré $\frac{d(d-1)}{2}$, degré maximal donné par les relations de Plücker.

Un des intérêts de l'image numérique est son bon comportement aux perturbations. Si A et B sont deux opérateurs bornés sur un espace de Hilbert, la distance de Hausdorff $d_H(W(A), W(B))$ est majorée par $\|A - B\|$. Pour les problèmes d'équations aux dérivées partielles, les approches variationnelles font intervenir des hypothèses sur les formes sesquilinéaires $\langle Au, u \rangle$, qui se traduisent assez bien en terme d'image numérique, (l'opérateur A y est généralement non borné). Les approximations numériques (méthodes d'éléments finis, méthodes dites spectrales, ondelettes, ...) utilisent des formes sesquilinéaires approchées $\langle A_h u_h, u_h \rangle$, l'image numérique de A_h hérite alors naturellement de propriétés analogues à celles de A .

Je voudrais cependant mettre un bémol à l'intérêt de cet outil. J'ai dit dans l'introduction que la majoration (1) pouvait, pour des opérateurs généraux, permettre à l'image numérique de jouer un rôle analogue au spectre pour les opérateurs normaux. Cependant on n'est pas dans une situation aussi favorable ; l'estimation (1) n'a pas la même précision que la théorie spectrale qui découple tous les modes. La convexité de $W(A)$ (qui contient $\sigma(A)$) peut rendre l'image numérique beaucoup trop grande. On peut par exemple construire un opérateur normal non borné pour lequel $W(A) = \mathbb{C}$ (on prend par exemple $A = \text{diag}(i e^i, 2i e^{2i}, \dots, ki e^{ki}, \dots)$ sur ℓ^2). Mais structurellement on ne peut pas espérer non plus pour les opérateurs généraux des outils aussi puissants que ceux permis par les opérateurs normaux...

5 Quelques raisons de croire en ma conjecture

La principale raison est la beauté du résultat que l'on obtiendrait si elle s'avérait exacte. Pour moi la beauté tient en la conjonction de la simplicité de l'énoncé et de la puissance des applications qui en découlent ; de plus celles-ci appartiennent à des secteurs assez divers des mathématiques. Il était donc naturel d'énoncer cette conjecture, ce que j'ai fait en 2002 ; depuis j'ai essayé en vain de lui trouver un contre-exemple.

On a $\mathcal{Q} = \sup_{\Omega} C(\Omega)$, Ω parcourant les ouverts convexes non vides. La constante $C(\Omega)$ étant invariante par action du groupe des similitudes, il est naturel d'espérer que la borne supérieure soit atteinte pour un Ω parfaitement symétrique, c'est à dire pour le cas du disque. Or dans ce cas on a $C(\Omega) = 2$. Il est aussi naturel de penser que ce maximum pourrait être atteint pour la situation limite où Ω est très plat, ce qui fait penser à la situation de la bande S_0 . Pour cette bande j'ai pu montrer que $C(S_0, 2) = 1,58766\dots$ et calculer numériquement des valeurs approchées $C(S_0, 4) = 1,672\dots$, $C(S_0, 6) = 1,726\dots$, $C(S_0, 8) = 1,765\dots$, ... (Par extrapolation empirique, cela pourrait donner $C(S_0) = 2$). Malheureusement la complexité des calculs augmente monstrueusement avec la dimension...

J'ai réussi à montrer que $\mathcal{Q}(2) = \mathcal{Q}_{cb}(2) = 2$. Pour $d = 3$ les essais numériques semblent montrer que les matrices A pour lesquels $W(A)$ est un disque (ce qui entraîne $\psi(A) \leq 2$) et $\psi(A) = 2$, correspondent à des maxima locaux de $\psi(A)$ (recherches aléatoires au voisinage de ces matrices). Plus généralement, je n'ai pas réussi à exhiber une matrice 3×3 telle que $\psi(A) > 2$.

6 Les essais numériques que j'ai effectués pour la bande et le secteur

Une difficulté importante pour un calcul numérique de $C(\Omega, d)$ est la prise en compte de la contrainte $W(A) \subset \Omega$. Cela se fait cependant assez simplement dans le cas de la bande ou du secteur. Nous regardons d'abord le cas de la bande $\Omega = S_0 := \{z \in \mathbb{C} ; |\text{Im } z| < 1\}$. Il est montré dans [3] qu'il existe une matrice $A \in \mathbb{C}^{d,d}$ et une fonction f holomorphe dans S_0 telle que

$$C(S_0, d) = \|f(A)\|, \quad \text{avec } W(A) \subset \overline{S_0} \text{ et } |f(z)| \leq 1 \text{ dans } S_0. \quad (3)$$

De plus la fonction f est de la forme

$$f(z) = \prod_{j=1}^{d-1} \frac{\text{th } \frac{\pi}{4} z - \zeta_j}{1 - \bar{\zeta}_j \text{th } \frac{\pi}{4} z} \quad \text{avec } |\zeta_j| \leq 1, \quad \text{ou encore } f(z) = \prod_{j=1}^{d-1} \frac{\exp \frac{\pi}{2} z - \gamma_j}{\exp \frac{\pi}{2} z + \bar{\gamma}_j} \quad \text{avec } \text{Re } \gamma_j \geq 0.$$

Notons que l'application conforme $z \mapsto \operatorname{th} \frac{\pi}{4} z$ envoie bijectivement la bande S_0 sur le disque unité D et l'application $z \mapsto \exp \frac{\pi}{2} z$ la bande S_0 sur le demi-plan $\operatorname{Re} z > 0$.

Puisque $W(A)$ et $\|f(A)\|$ sont invariants lorsque l'on remplace A par U^*AU , avec U unitaire, on peut imposer que la matrice A qui réalise (3) soit de la forme $A = B + iC$, avec B auto-adjointe et C diagonale réelle. La condition $W(A) \subset \overline{S_0}$ s'écrit alors simplement

$$A = B + iC, \quad \text{avec } B = B^*, \quad C = \operatorname{diag}(c_i), \quad c_i \in [-1, 1], i = 1, \dots, d.$$

En fait on peut imposer que l'on ait $c_i = \pm 1$. En effet, si on suppose que pour un certain k on a $|c_k| < 1$, on a alors, pour tout z vérifiant $|z| \leq 1 - |c_k|$, $W(A + zE_k) \subset \overline{S_0}$, en notant par E_k la matrice $d \times d$ dont le seul élément non nul vaut 1, et se trouve en position (k, k) . On a alors $\|f(A + zE_k)\| \leq C(S_0, d) = \|f(A)\|$. Par le principe du maximum pour la fonction holomorphe $f(A + zE_k)$, on en déduit que $\|f(A + zE_k)\| = \|f(A)\|$, pour $|z| \leq 1 - |c_k|$. Il en résulte en particulier que l'on peut remplacer la valeur de c_k par 1 si $c_k \geq 0$, ou par -1 sinon, sans changer la valeur de $\|f(A)\|$.

On en déduit que l'on peut se limiter à regarder des matrices A de la forme

$$A = \begin{pmatrix} D_1 & E \\ E^* & D_2 \end{pmatrix} + i \begin{pmatrix} I_k & 0 \\ 0 & -I_{d-k} \end{pmatrix}, \quad \text{avec } 1 \leq k < d. \quad (4)$$

(On ne peut pas avoir $k = 0$, ni $k = d$, sinon autrement la matrice A serait normale, ce qui serait incompatible avec $C(S_0, d) > 1$.) En utilisant l'invariance par similitude unitaire, on peut supposer que D_1 et D_2 sont des matrices diagonales réelles, et de plus que les éléments de la première ligne et de la dernière colonne de E sont réels. L'invariance par translation horizontale, permet de supposer que $\operatorname{trace}(A) = 0$; quitte à remplacer A par $-A$, puis à effectuer une permutation des blocs, on peut supposer aussi que $k \leq d/2$. On a donc

$$C(S_0, d) = \max_{1 \leq k \leq d/2} \max_{D_1, D_2, E, \gamma} \|f(A)\|, \quad \text{avec } f(z) = \prod_{j=1}^{d-1} \frac{\exp \frac{\pi}{2} z - \gamma_j}{\exp \frac{\pi}{2} z + \bar{\gamma}_j}.$$

On est donc ramené à un problème d'optimisation portant, pour chaque valeur de k , sur $2(k(d-k) + d - 1)$ variables, dont $(d - 1)$ avec la contrainte de positivité ($\operatorname{Re} \gamma_i \geq 0$). Pour $d = 2$, $k = 1$, on a un problème d'optimisation à 4 variables, mais pour $d = 4$ et $k = 2$, déjà 14 variables...

Pour $d = 2$, on retrouve la valeur $C(S_0, 2) = 1.5876598\dots$ (montrée rigoureusement dans [3]) sans aucune difficulté.

Pour $d = 4$, le programme d'optimisation (avec une donnée initiale aléatoire) a tendance à converger à peu près une fois sur deux vers le maximum local $C(S_0, 2)$, à peu près une fois sur deux vers un maximum local $1.59400\dots$, mais aussi, près d'une fois sur dix vers la valeur 1.6723401 . Je crois que cette dernière correspond à la valeur de $C(S_0, 4)$. On remarque alors que la matrice A correspondante possède de nombreuses propriétés de symétrie

$$A = \begin{pmatrix} D + iI & E \\ E & D - iI \end{pmatrix}, \quad \text{avec } D = \begin{pmatrix} x_1 & 0 \\ 0 & -x_1 \end{pmatrix}, \quad E = \begin{pmatrix} x_2 & x_3 \\ x_3 & x_2 \end{pmatrix},$$

$$x_1 = 2.3816\dots, \quad x_2 = 1.388\dots, \quad x_3 = 1.2523\dots, \quad \gamma_1 = 8.566\dots, \quad \gamma_2 = 1, \quad x_3 = 1/\gamma_3.$$

Un problème ouvert est de montrer que l'on peut se limiter à regarder les matrices possédant ces symétries. Cela permet alors d'envisager des calculs en dimensions plus élevées. J'ai ainsi obtenu la valeur $1.72662\dots$ pour $d = 6$ et 1.764577 pour $d = 8$.

Essais pour le secteur $S_\alpha = \{z \in \mathbb{C}; z \neq 0 \text{ et } |\arg z| < \alpha\}$, $0 < \alpha < \pi/2$. On vérifie facilement que la condition $W(A) \subset \overline{S_\alpha}$ est équivalente à la possibilité d'écrire A sous la forme $A = B(\cos \alpha I + i \sin \alpha C)B$, avec B et C auto-adjointes et $\|C\| \leq 1$. Il est montré dans [3] qu'il existe une matrice $A \in \mathbb{C}^{d,d}$ et une fonction f holomorphe dans S_α telle que

$$C(S_\alpha, d) = \|f(A)\|, \quad \text{avec } W(A) \subset \overline{S_\alpha}, \quad |f(z)| \leq 1 \text{ dans } S_\alpha, \quad (5)$$

la fonction f étant de la forme

$$f(z) = \prod_{j=1}^{d-1} \frac{z^s - \gamma_j}{z^s + \bar{\gamma}_j}, \quad \text{avec } s = \frac{\pi}{2\alpha} \text{ et } \operatorname{Re} \gamma_j \geq 0.$$

Avec le même raisonnement que pour la bande, on peut supposer que la matrice C est diagonale et que ses valeurs propres sont $+1$ ou -1 . La matrice A est alors de la forme

$$A = \begin{pmatrix} D_1 & E \\ E^* & D_2 \end{pmatrix} \begin{pmatrix} e^{i\alpha} I_k & 0 \\ 0 & e^{-i\alpha} I_{d-k} \end{pmatrix} \begin{pmatrix} D_1 & E \\ E^* & D_2 \end{pmatrix}, \quad \text{avec } 1 \leq k < d. \quad (6)$$

et on peut supposer les matrices D_1 et D_2 diagonales réelles. On se retrouve alors dans une situation similaire à celle de la bande pour lancer un programme d'optimisation basé sur la formule

$$C(S_\alpha, d) = \max_{1 \leq k \leq d/2} \max_{D_1, D_2, E, \gamma} \|f(A)\|, \quad \text{avec } f(z) = \prod_{j=1}^{d-1} \frac{z^s - \gamma_j}{z^s + \bar{\gamma}_j} \quad \text{et } s = \frac{\pi}{2\alpha}.$$

Les essais numériques s'avèrent un peu plus délicats que pour la bande. Pour $d = 4$, très souvent le programme converge vers des maxima non globaux, ou s'arrête par génération de INF ou de NAN (cela est vrai surtout lorsque α est petit, donc $s = \frac{\pi}{2\alpha}$ grand, les calculs de z^s engendrent des instabilités. D'autre part, si on converge vers un maximum local correspondant à $d = 2$, certains γ_i peuvent tendre vers 0 ou vers l'infini...). J'ai pu suivre par continuation un maximum local de $\|f(A)\|$ convergeant vers 1.587... (celui correspondant à $C(S_\alpha, 2)$) et un autre convergeant vers 1.672... Les valeurs se croisent vers $\alpha = 2\pi/13$.

Mes essais numériques laissent penser que, pour le quart de plan, $C(S_{\pi/4}, 4) = C(S_{\pi/4}, 2) = \sqrt{2}$. Est-ce que cela reste vrai pour tout d ?

Le quart de plan correspond à une géométrie simple, la transformation conforme $z \mapsto z^2$ de $S_{\pi/4}$ dans le demi-plan $\operatorname{Re} z > 0$ est elle aussi très simple. Mais je n'ai pas réussi à en profiter pour obtenir un résultat meilleur dans ce cas.

7 Essais numériques sur les matrices 3×3

Ces essais sont basés sur la formule

$$Q(3) = \max_{A \in \mathbb{C}^{3,3}} \psi(A).$$

Le problème est alors le calcul de $\psi(A)$.

Remarque. En utilisant que $\psi(A) = \psi(U^*AU)$, si U est unitaire, et $\psi(A) = \psi(\lambda A + \mu I)$, si $\lambda \neq 0$, on peut se limiter à ne considérer que des matrices A triangulaires supérieures, à trace nulle, à éléments non diagonaux réels, positifs ou nuls, tels que $\sum_{j>i} a_{ij}^2 = 1$. La matrice A ne dépend plus alors que de 6 paramètres réels.

Soit A une matrice n'ayant aucune valeur propre sur la frontière de $W(A)$. L'intérieur de $W(A)$ n'est alors pas vide, et on peut considérer une application conforme a qui l'envoie sur le disque unité. On a alors $\psi(A) = \psi_D(a(A))$. Cela permet de décomposer le calcul de $\psi(A)$ en trois étapes :

Étape 1. Calcul de la frontière de $W(A)$.

Étape 2. Calcul de la matrice $B = a(A)$.

Étape 3. Calcul de $\psi_D(B)$.

Je vais détailler comment j'ai effectué chacune d'elles.

Étape 1. J'ai choisi des valeurs équidistantes pour les angles $\theta_j = \frac{2j\pi}{2n+1}$, $j = 0, 1, \dots, 2n$. Ensuite, pour chaque valeur de j , j'ai cherché le point de $z_j \in \partial W(A)$ pour lequel la normale extérieure a pour composantes $(\cos \theta_j, \sin \theta_j)$. Il est obtenu en prenant $z_j = w_j^* A w_j / w_j^* w_j$, où w_j est un vecteur propre correspondant à la plus grande valeur propre de la matrice $\cos \theta_j B + \sin \theta_j C$ (avec la notation $A = B + iC$).

Remarque. Si l'on suppose que les valeurs propres de A sont intérieures à son image numérique (ce qui se voit facilement si A est sous forme triangulaire), génériquement la plus grande valeur propre de $\cos \theta_j B + \sin \theta_j C$ est simple pour tout θ , et la frontière de $W(A)$ est analytique. Hors de cette situation générique, il y a au plus une valeur de θ pour laquelle cette plus grande valeur propre est multiple, elle est alors double, la frontière est le raccordement de classe C^1 d'une cardioïde avec un segment de droite.

Étape 2. Pour le calcul de $B = a(A)$, on utilise la formule des différences divisées de Newton,

$$B = a(\lambda_1)I + a[\lambda_1, \lambda_2](A - \lambda_1 I) + a[\lambda_1, \lambda_2, \lambda_3](A - \lambda_1 I)(A - \lambda_2 I),$$

où les λ_j sont les valeurs propres de A . pour cela il suffit donc de calculer les différences divisées correspondantes, ce qui nécessite au préalable de connaître la transformation conforme a . On se ramène au cas où 0 est intérieur à l'image numérique (par exemple en choisissant d'avoir $\text{trace}(A) = 0$). Pour définir a de manière unique (au produit par un nombre complexe de module 1 près), on choisira alors une transformation conforme telle que $a(0) = 0$. On peut alors l'écrire sous la forme $a(z) = z \exp(u+iv)$, les fonctions $u(z)$ et $v(z)$ étant harmoniques, à valeurs réelles. Remarquons que l'on a $u(z) = -\log |z|$ sur $\partial W(A)$, ce qui détermine entièrement u harmonique dans $W(A)$.

Nous supposons maintenant que la frontière de $W(A)$ est connue par une paramétrisation régulière $\partial W(A) = \{\sigma(\theta); \theta \in [0, 2\pi]$. On sait alors qu'il existe une fonction $q(\cdot)$ unique, 2π périodique, à valeurs réelles, telles que, pour tout z intérieur à $W(A)$,

$$(u+iv)(z) = \int_0^{2\pi} q(\theta) \log(\sigma(\theta) - z) d\theta, \quad \text{et} \quad \int_0^{2\pi} q(\theta) d\theta = 0.$$

Pour obtenir cette fonction q , on regarde la partie réelle de l'équation précédente aux points $z = \sigma(\varphi) \in \partial W(A)$, ce qui nous donne

$$\int_0^{2\pi} q(\theta) \log |\sigma(\theta) - \sigma(\varphi)| d\theta = -\log |\sigma(\varphi)|, \quad \text{pour tout } \varphi \in [0, 2\pi[,$$

ce qui s'écrit encore

$$\int_0^{2\pi} q(\theta) \log \left| \frac{\sigma(\theta) - \sigma(\varphi)}{e^{i\theta} - e^{i\varphi}} \right| d\theta + \int_0^{2\pi} q(\theta) \log |e^{i\theta} - e^{i\varphi}| d\theta = -\log |\sigma(\varphi)|, \quad \forall \varphi \in [0, 2\pi[.$$

J'ai discrétisé cette équation en utilisant la paramétrisation $\sigma(\theta)$ de $\partial W(A)$ obtenue à l'étape 1, en approchant $q(\cdot)$ par un polynôme trigonométrique $q_n(\cdot)$ de degré n et en utilisant une méthode

de collocation aux points θ_j , $j = 0, 1, \dots, 2n$ (c'est pour cela que j'avais choisi un nombre impair de points à l'étape 1, il est bien connu que cela est nécessaire pour cette méthode de collocation). On obtient ainsi une approximation $q_j = q_n(\theta_j)$ en résolvant le système

$$\frac{2\pi}{2n+1} \sum_{j=0}^{2n} q_j \log \left| \frac{\sigma(\theta_j) - \sigma(\theta_i)}{e^{i\theta_j} - e^{i\theta_i}} \right| + \int_0^{2\pi} q_n(\theta) \log |e^{i\theta} - e^{i\theta_i}| d\theta = -\log |\sigma(\theta_i)|,$$

pour $i = 0, 1, \dots, 2n$.

La première intégrale a été approchée par la formule du trapèze ; bien sûr, pour $j = i$ on doit remplacer $\log \left| \frac{\sigma(\theta_j) - \sigma(\theta_i)}{e^{i\theta_j} - e^{i\theta_i}} \right|$ par $\log |\sigma'(\theta_i)|$. Rappelons que pour l'intégrale restante on a la formule

$$\int_0^{2\pi} q_n(\theta) \log |e^{i\theta} - e^{i\theta_i}| d\theta = -\frac{2\pi}{2n+1} \sum_{j=0}^{2n} c(j-i) q_j,$$

avec $c(k) = c(-k) = \sum_{j=0}^n \frac{\cos j\theta_k}{j}$.

Tant que la frontière de l'image numérique est analytique, cette méthode est très efficace (convergence exponentielle par rapport à n). Malheureusement (et c'est bien compréhensible) elle se dégrade très fortement lorsque l'on approche la situation non générique décrite dans la remarque de l'étape 1.

Étape 3. Pour le calcul de $\psi_D(B)$ j'ai utilisé un programme d'optimisation à partir de la caractérisation

$$\psi_D(B) := \sup_{\zeta_1, \zeta_2} \{ \|g(B)\| ; g(z) = \frac{z - \zeta_1}{1 - \bar{\zeta}_1 z} \frac{z - \zeta_2}{1 - \bar{\zeta}_2 z}, \quad \zeta_1, \zeta_2 \in D \}.$$

Mon programme s'avère fiable et précis pour le calcul de $\psi(A)$ lorsque l'image numérique a une frontière analytique, mais sa dégradation (par instabilité) lorsque l'on approche la situation enveloppe convexe d'une cardioïde, fait qu'on ne peut pas l'inclure dans un programme d'optimisation pour calculer

$$\mathcal{Q}(3) = \max_A \psi(A).$$

Mes essais m'ont convaincu que $\mathcal{Q}(3) = 2$. La méthode décrite ci-dessus se généralise assez immédiatement aux dimensions supérieures, mais je ne l'ai pas encore fait...

8 Quelques compléments sur les constantes associées à un domaine convexe

Tout d'abord, je vais donner un résultat (sans doute bien connu) concernant le lien entre la convexité et les représentations conformes de l'ouvert correspondant

Lemme 1. *Soit b une application conforme bijective du disque unité ouvert D sur un ouvert Ω . Alors les 3 propriétés suivantes sont équivalentes :*

- i) Ω est convexe,
- ii) $\forall \mu > 0, \nu > 0$ avec $\mu + \nu < 2\pi$, $\forall 0 \neq z \in D$, on a $0 < \arg \left(\frac{b(e^{i\mu} z) - b(z)}{b(z) - b(e^{-i\nu} z)} \right) < \pi$,
- iii) $\forall z \in D$, on a $\operatorname{Re} \left(1 + \frac{z b''(z)}{b'(z)} \right) > 0$.

Démonstration. $i) \implies ii)$. Supposons Ω convexe. La fonction b se prolonge alors continûment et biunivoquement jusqu'à la frontière ∂D (avec un point envoyé à l'infini si Ω est non borné).

Soit μ et ν fixés, on pose, pour $z \neq 0$, $\varphi(z) = \arg\left(\frac{b(e^{i\mu}z) - b(z)}{b(z) - b(e^{-i\nu}z)}\right)$. On a clairement $0 \leq \varphi(z) \leq \pi$, pour tout $z \in \partial D$, par convexité de la frontière et conservation de l'orientation. La fonction φ est harmonique, elle se prolonge continûment en 0 par $\varphi(0) = \arg\left(\frac{e^{i\mu}-1}{1-e^{-i\nu}}\right) = \frac{\mu+\nu}{2}$. On déduit du principe du maximum que l'on a $0 < \varphi(z) < \pi$, pour tout $z \in D$.

$ii) \implies iii)$. Soit $0 \neq z \in D$. On a nécessairement $b'(z) \neq 0$ puisque b est bijectif. Pour $\mu = \nu > 0$ voisin de 0, on écrit

$$\begin{aligned} \frac{b(e^{i\mu}z) - b(z)}{b(z) - b(e^{-i\mu}z)} &= 1 + \frac{b(e^{i\mu}z) - 2b(z) + b(e^{-i\mu}z)}{b(z) - b(e^{-i\mu}z)} \\ &= 1 + i\mu \operatorname{Re}\left(1 + \frac{z b''(z)}{b'(z)}\right) + O(\mu^2). \end{aligned}$$

On a donc

$$\arg\left(\frac{b(e^{i\mu}z) - b(z)}{b(z) - b(e^{-i\mu}z)}\right) = \mu\left(1 + \operatorname{Re}\frac{z b''(z)}{b'(z)}\right) + O(\mu^2).$$

Cela montre bien que $1 + \operatorname{Re}\frac{z b''(z)}{b'(z)} \geq 0$ (cela reste vrai de manière évidente, avec une inégalité stricte, pour $z = 0$). Par le principe du maximum, on a la stricte positivité pour tout $z \in D$.

$iii) \implies i)$. Soit Γ_r l'image par b du cercle $C_r := \{z = r e^{i\theta}; \theta \in [0, 2\pi]\}$. La courbure $c(z)$ de Γ_r au point $b(z)$, $z \in C_r$ est donnée par la formule

$$c(z) = \frac{1}{r |b'(z)|} \operatorname{Re}\left(1 + \frac{z b''(z)}{b'(z)}\right).$$

Il en résulte que l'ouvert délimité par Γ_r est strictement convexe pour tout $r \in]0, 1[$. Par passage à la limite quand $r \rightarrow 1$, on obtient la convexité de Ω . \square

Remarque. Il résulte de la démonstration que tout domaine convexe borné peut être approché d'aussi près que l'on veut (au sens de Hausdorff) par un domaine ayant une frontière analytique avec une courbure strictement positive en tout point.

Lemme 2. Soit $\Omega \subset \mathbb{C}$ un domaine convexe qui n'est, ni \mathbb{C} , ni un demi-plan. On note a une transformation conforme de Ω sur le disque unité D . On a alors

$$C(\Omega, 2) \geq 2 \sup_{z_1 \in \Omega} \frac{d(z_1, \partial\Omega) |a'(z_1)|}{1 - |a(z_1)|^2} > 1.$$

Démonstration. Soit $z_1 \in \Omega$ fixé; il existe alors $z_0 \in \partial\Omega$ tel que $\gamma = |z_1 - z_0| = d(z_1, \partial\Omega)$ soit la distance de z_1 à la frontière de Ω . On considère alors la matrice A et la fonction f définies par

$$A = \begin{pmatrix} z_1 & 2\gamma \\ 0 & z_1 \end{pmatrix}, \quad f(z) = \frac{a(z) - a(z_1)}{1 - \overline{a(z_1)}a(z)}, \quad \text{d'où } f(A) = 2\gamma f'(z_1) \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

L'image numérique de A est le cercle centré en z_1 de rayon γ ; on a donc $W(A) \subset \overline{\Omega}$. Par ailleurs la fonction f est de module majoré par 1 sur Ω (c'est en fait une représentation conforme de Ω sur D). On a donc

$$C(\Omega, 2) \geq \|f(A)\| = 2\gamma |f'(z_1)| = \frac{d(z_1, \partial\Omega) |a'(z_1)|}{1 - |a(z_1)|^2}.$$

Considérons maintenant les fonctions

$$c(z) = \frac{f(z_0) + f(z)}{f(z_0) - f(z)}, \quad \text{et} \quad \varphi(z) = (z - z_0)c'(z) + c(z).$$

Notons que c envoie biunivoquement Ω sur le demi-plan $\operatorname{Re} z > 0$, et la frontière de Ω sur l'axe imaginaire. On remarque aussi que $|c'(z_1)| = 2|f'(z_1)|$. Pour montrer le lemme il suffit donc de montrer que $|z_1 - z_0| |c'(z_1)| > 1$.

Supposons tout d'abord la frontière de Ω analytique. La fonction c est alors de la forme $c(z) = \frac{\alpha}{z - z_0} + \beta(z)$, avec β analytique dans un voisinage de $\overline{\Omega}$. La fonction φ est donc holomorphe dans ce voisinage. On remarque par ailleurs que, si $z \in \partial\Omega$, $z \neq z_0$, on a $\operatorname{Re} \varphi(z) = \operatorname{Re} ((z - z_0)c'(z)) \leq 0$ et, puisque Ω n'est pas un demi-plan, $\operatorname{Re} \varphi$ n'est pas identiquement nul. On a donc $\operatorname{Re} \varphi < 0$ dans Ω par le principe du maximum. En particulier au point z_1 on a $\operatorname{Re}(z_1 - z_0)c'(z_1) + 1 = \operatorname{Re} \varphi(z_1) < 0$. Cela implique $|(z_1 - z_0)c'(z_1)| > 1$.

Lorsque la frontière de Ω n'est pas analytique, on peut approcher Ω par un Ω_n analytique, avec $z_0 \in \partial\Omega \cap \partial\Omega_n$, trouver une transformation conforme a_n qui converge vers a dans Ω . On obtient alors par passage à la limite, $\operatorname{Re} \varphi \leq 0$ dans Ω . Si on avait $\operatorname{Re} \varphi = 0$ en un point de Ω donc en tout point, on aurait alors $\varphi = ia$ (constant) dans Ω , d'où l'on déduirait que Ω serait un demi-plan. On a donc $\operatorname{Re} \varphi < 0$ dans Ω et on conclut comme précédemment. \square

Remarque. En utilisant la matrice A et la fonction f définies par

$$A = \begin{pmatrix} 0 & 2r \\ 0 & 0 \end{pmatrix}, \quad f(z) = z,$$

on obtient aussi la minoration plus brutale

$$C(\Omega, 2) \geq \frac{2r}{R},$$

sous l'hypothèse que le domaine convexe Ω contienne la boule de centre 0 et de rayon r et soit contenu dans la boule de centre 0 et de rayon R .

Lemme 3. *Soit Ω un domaine convexe dont la frontière présente un point anguleux d'angle 2α . Alors on a $C(\Omega, d) \geq C(S_\alpha, d)$.*

Démonstration. Sans perte de généralité, on peut supposer que le point anguleux est l'origine, et que Ω est contenu dans le secteur $S_\alpha = \{z \in \mathbb{C}; z \neq 0 \text{ et } |\arg z| < \alpha\}$, avec tangence à l'origine. Soit $\gamma < C(S_\alpha, d)$; il existe alors une matrice $A \in \mathbb{C}^{d,d}$, avec $W(A) \subset S_\alpha$, et une fraction rationnelle r bornée par 1 dans S_α , telle que $\|r(A)\| \geq \gamma$. Notons que, quel que soit $\varepsilon > 0$, si on pose $r_\varepsilon(z) = r(z/\varepsilon)$, on a encore $W(\varepsilon A) \subset S_\alpha$ et r_ε est aussi borné par 1 dans S_α . Mais pour ε assez petit on a $W(\varepsilon A) \subset \Omega$; comme r_ε est a fortiori borné par 1 dans Ω , on en déduit $\gamma \leq \|r(A)\| = \|r_\varepsilon(\varepsilon A)\| \leq C(\Omega, d)$. \square

References

- [1] C. Badea, M. Crouzeix, B. Delyon, Convex domains and K-spectral sets, *Math. Z.* **252**, no. 2, 345–365, 2006.
- [2] B. Beckermann, M. Crouzeix, Operators with numerical range in a conic domain, preprint.
- [3] M. Crouzeix, Bounds for analytic functions of matrices, *Int. Equ. Op. Th.* **48**, 461–477, 2004.

- [4] M. Crouzeix, Numerical range and functional calculus in Hilbert space, preprint, available from <http://perso.univ-rennes1.fr/michel.crouzeix/>
- [5] M. Crouzeix, B. Delyon, Some estimates for analytic functions of strip or sectorial operators, *Arch. Math.*, **81** (2003), 553-566.
- [6] M. Crouzeix, Operators with numerical range in a parabola, *Arch. Math.*, **82** (2004), 517-527.
- [7] B. & F. Delyon, Generalization of Von Neumann's spectral sets and integral representation of operators, *Bull. Soc. Math. France*, **1**, (1999), 25-42.
- [8] K. E. Gustafson, D. K. M. Rao, *Numerical Range*, Universitext, Springer-Verlag, 1997.
- [9] V. Paulsen, *Completely bounded maps and operator algebras*, Cambridge Univ. Press, 2002.
- [10] V. Paulsen, K-spectral values for finite matrices, *J. Operator Theory*, vol. **18**, no 2, 249-263, 1987.
- [11] G. Pisier, A polynomially bounded operator on a Hilbert space which is not similar to a contraction, *J. Amer. Math. Soc.* **10** (1997), 351-369.
- [12] J. von Neumann, Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes, *Math. Nachrichten* **4** (1951) 258-281.
- [13] R. Smith Completely bounded maps between C^* algebras, *J. London Math. Soc.* **27** (1983) 157-166.

Institut de Recherche Mathématique de Rennes, UMR CNRS n° 6625
 Université de Rennes 1, Campus de Beaulieu, 35042 RENNES Cedex, France
 michel.crouzeix@univ-rennes1.fr