

Sur une constante liée aux matrices 3×3

Michel CROUZEIX

1 Introduction.

Soit $A \in \mathbb{C}^{d,d}$ une matrice carrée à coefficients complexes. On lui associe son image numérique $W(A)$ définie par $W(A) := \{v^*Av; v^*v = 1, v \in \mathbb{C}^d\}$. On définit une constante \mathcal{Q}_d par

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

On sait que $\mathcal{Q}_2 = 2$ [1], de plus nous avons la borne uniforme $\mathcal{Q}_d < 11.08$, voir [2]. Dans ce manuscrit nous nous intéressons plus précisément à la constante \mathcal{Q}_3 . Pour cela il nous sera utile d'introduire les fonctions

$$\begin{aligned} \psi(A) &= \sup_p \{\|p(A)\|; |p| \leq 1 \text{ dans } W(A), p \text{ polynôme}\}, & A \in \mathbb{C}^{d,d}, \\ \psi_D(M) &= \sup_f \{\|f(M)\|; |f| \leq 1 \text{ dans } D, f \in \mathcal{H}^\infty(D)\}, & M \in \mathbb{C}^{d,d}, \sigma(M) \subset D. \end{aligned}$$

On a noté ici $D := \{z \in \mathbb{C}; |z| < 1\}$ le disque ouvert unité du plan complexe et $\mathcal{H}^\infty(D)$ l'espace de Hardy des fonctions holomorphes et bornées sur D . La condition $\sigma(M) \subset D$ implique que $f(M)$ est bien défini (par exemple par la formule de Cauchy) dès que $f \in \mathcal{H}^\infty(D)$.

Remarque. Si on choisit

$$A_1 = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{ou encore } A_2 = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix},$$

on a, par un calcul simple, $W(A_1) = W(A_2) = \overline{D}$ et $\|A_1\| = 2 = \|A_2\|$, donc $\psi(A_1) \geq 2$ et de même $\psi(A_2) \geq 2$ (en fait on peut montrer que $\psi(A_1) = \psi(A_2) = 2$). Par suite $\mathcal{Q}_3 \geq 2$.

Il y a donc deux possibilités

- $\mathcal{Q}_3 = 2$, ce qui est ma conviction, la constante \mathcal{Q}_3 est donc atteinte pour des matrices A dont l'image numérique est un disque. Nous n'étudierons pas plus en détail ce cas, mais il resterait la question ouverte : est-ce que $\psi(A) = 2$ implique $W(A)$ est un disque ? (La réponse est positive pour les matrices 2×2 , c.f. [1]).
- Le reste de ce papier est donc écrit sous la condition hypothétique $\mathcal{Q}_3 > 2$.

Dans le reste de ce papier, nous supposons donc $\mathcal{Q}_3 > 2$ et nous regardons essentiellement des matrices 3×3 .

2 Étude de $\psi_D(M)$ et de $\psi(A)$.

Soit $M \in \mathbb{C}^{3,3}$ et soit $\lambda_1, \lambda_2, \lambda_3$ ses valeurs propres. On suppose qu'elles appartiennent à D . Rappelons que, pour f holomorphe dans D , on a la représentation de Newton

$$f(M) = f[\lambda_1]I + f[\lambda_1, \lambda_2](M - \lambda_1 I) + f[\lambda_1, \lambda_2, \lambda_3](M - \lambda_1 I)(M - \lambda_2 I),$$

avec les différences divisées

$$\begin{aligned} f[\lambda_1] &= f(\lambda_1), & f[\lambda_1, \lambda_2] &= \int_0^1 f'((1-t)\lambda_1 + t\lambda_2) dt, \\ f[\lambda_1, \lambda_2, \lambda_3] &= \int_0^1 \int_0^1 (1-t)f''((1-t)(1-s)\lambda_1 + (1-t)s\lambda_3 + t\lambda_2) ds dt. \end{aligned}$$

Pour des valeurs propres distinctes, on a les expressions plus simples

$$f[\lambda_1, \lambda_2] = \frac{f(\lambda_1) - f(\lambda_2)}{\lambda_1 - \lambda_2}, \text{ si } \lambda_1 \neq \lambda_2, \quad \text{et } f[\lambda_1, \lambda_2, \lambda_3] = \frac{f[\lambda_2, \lambda_3] - f[\lambda_1, \lambda_2]}{\lambda_3 - \lambda_1}, \text{ si } \lambda_3 \neq \lambda_1.$$

Soit f une fonction holomorphe donnée, satisfaisant $\|f\|_\infty := \sup_{z \in D} |f(z)| = 1$. On lui associe la fonction holomorphe g qui réalise le minimum de $\|g\|_\infty$ sous les contraintes

$$g[\lambda_1] = f[\lambda_1], \quad g[\lambda_1, \lambda_2] = f[\lambda_1, \lambda_2] \quad \text{et} \quad g[\lambda_1, \lambda_2, \lambda_3] = f[\lambda_1, \lambda_2, \lambda_3].$$

D'après l'annexe, une telle fonction existe, est unique, et est de la forme

$$g(z) = m e^{i\theta} b(z) = m e^{i\theta} \frac{z - \alpha_1}{1 - \bar{\alpha}_1 z} \frac{z - \alpha_2}{1 - \bar{\alpha}_2 z}, \quad \alpha_1, \alpha_2 \in \bar{D}, \quad m > 0, \quad \theta \in \mathbb{R}.$$

De plus, on a $m = \|g\|_\infty \leq 1$ et $f(M) = g(M) = m e^{i\theta} b(M)$, d'où $\|b(M)\| \geq \|f(M)\|$.

Il en résulte que l'on a

$$\psi_D(M) = \max\{\|b(M)\|; b(z) = \frac{z - \alpha_1}{1 - \bar{\alpha}_1 z} \frac{z - \alpha_2}{1 - \bar{\alpha}_2 z}, \quad \alpha_1, \alpha_2 \in \bar{D}\}. \quad (1)$$

(La fonction $(\alpha_1, \alpha_2) \mapsto \|b(M)\|$ étant continue sur le compact $\bar{D} \times \bar{D}$, le maximum est bien atteint et est fini.)

Soit maintenant une matrice A ; on suppose que son spectre $\sigma(A)$ est contenu dans l'intérieur de l'image numérique $W(A)$. D'après les résultats sur les transformations conformes, il existe une bijection bicontinue a de l'ensemble convexe $W(A)$ sur \bar{D} qui est holomorphe à l'intérieur de $W(A)$. Soit d l'application réciproque de a . Si on pose $M = a(A)$ et $f(z) = p(d(z))$, p étant un polynôme, on a $p(A) = f(M)$; de plus $f \in \mathcal{H}^\infty(D)$ et $\|f\|_\infty = \sup_{z \in W(A)} |p(z)|$. On en déduit que $\psi(A) \leq \psi_D(M)$. D'autre part, si b est de la forme donnée dans (1), et si on pose $g(z) = b(a(z))$, alors g est holomorphe à l'intérieur de $W(A)$ et continue dans $W(A)$. On a $g(A) = b(M)$ et de plus g est limite uniforme dans $W(A)$ de polynômes. Il en résulte que

$$\psi(A) = \max\{\|g(A)\|; g(z) = \frac{a(z) - \alpha_1}{1 - \bar{\alpha}_1 a(z)} \frac{a(z) - \alpha_2}{1 - \bar{\alpha}_2 a(z)}, \quad \alpha_1, \alpha_2 \in \bar{D}\}. \quad (2)$$

3 Quelques propriétés de l'image numérique.

Les propriétés suivantes de l'image numérique nous seront utiles et sont faciles à vérifier.

- $W(A)$ est compacte et contient le spectre $\sigma(A)$ de la matrice A ,
- $W(\lambda A + \mu I) = \lambda W(A) + \mu$,
- si U est une matrice unitaire alors $W(U^*AU) = W(A)$,
- si A et B sont deux matrices carrées de même dimension, alors la distance de Hausdorff entre leurs images numériques vérifie $d_H(W(A), W(B)) \leq \|A - B\|$.

Nous donnons maintenant une démonstration élémentaire du résultat de Toeplitz et Hausdorff

Lemme 1. *L'image numérique est convexe.*

Démonstration. Soit $\lambda = \frac{\langle Af, f \rangle}{\langle f, f \rangle}$ et $\mu = \frac{\langle Ag, g \rangle}{\langle g, g \rangle}$ deux points distincts de $W(A)$. Remplaçant si besoin g par $e^{i\theta}g$, on peut supposer que

$$\begin{aligned} \frac{\langle (A-\lambda)f, g \rangle}{\mu - \lambda} - \frac{\langle (A^* - \bar{\lambda})f, g \rangle}{\bar{\mu} - \bar{\lambda}} &\in \mathbb{R}, \quad \text{donc} \\ \frac{\langle (A-\lambda)f, g \rangle + \langle (A-\lambda)g, f \rangle}{\mu - \lambda} &= \frac{\langle (A-\lambda)f, g \rangle}{\mu - \lambda} - \frac{\langle (A^* - \bar{\lambda})f, g \rangle}{\bar{\mu} - \bar{\lambda}} + 2 \operatorname{Re} \frac{\langle (A-\lambda)g, f \rangle}{\mu - \lambda} \in \mathbb{R}. \end{aligned}$$

Posons alors, pour $t \in \mathbb{R}$,

$$\begin{aligned} \varphi(t) &:= \frac{\langle A(f+tg), f+tg \rangle - \lambda \langle f+tg, f+tg \rangle}{(\mu - \lambda) \langle f+tg, f+tg \rangle} \\ &= t \frac{\langle (A-\lambda)f, g \rangle + \langle (A-\lambda)g, f \rangle}{(\mu - \lambda) \langle f+tg, f+tg \rangle} + t^2 \frac{\langle g, g \rangle}{\langle f+tg, f+tg \rangle} \in \mathbb{R}. \end{aligned}$$

La fraction rationnelle φ est continue sur $\overline{\mathbb{R}}$ et prend ses valeurs dans \mathbb{R} . On a $\varphi(0) = 0$, $\varphi(\infty) = 1$ d'où, pour tout $\tau \in (0, 1)$, il existe $t \in \mathbb{R}$ avec $\varphi(t) = \tau$. En remarquant que

$$(1-\tau)\lambda + \tau\mu = (\lambda - \mu)\varphi(t) + \lambda = \frac{\langle A(f+tg), f+tg \rangle}{\langle f+tg, f+tg \rangle} \in W(A),$$

on en déduit que le segment joignant λ à μ est contenu dans $W(A)$. \square

Nous aurons aussi besoin des propriétés suivantes

Lemme 2. *a) Si $\lambda \in \sigma(A) \cap \partial W(A)$, alors $\mathbb{C}^d = \operatorname{Ker}(A - \lambda I) \oplus \operatorname{Im}(A - \lambda I)$. De plus cette décomposition est orthogonale.*

b) Si λ est un point anguleux de $\partial W(A)$, alors λ est valeur propre de A .

Démonstration. a) Quitte à utiliser la propriété $W(\mu A - \lambda I) = \mu W(A) - \lambda$, on peut se ramener au cas où $0 \in \sigma(A) \cap \partial W(A)$ et $\operatorname{Re}(Av, v) \geq 0$, pour tout $v \in \mathbb{C}^d$. On écrit alors $A = B + iC$, avec $B = B^*$, $C = C^*$. Si $u \in \operatorname{Ker} A$, on a clairement $(Bu, u) = \operatorname{Re}(Au, u) = 0$, or $(Bv, v) \geq 0$ pour tout v . On a donc $Bu = 0$. Par suite $\operatorname{Ker} A = \operatorname{Ker} A^* = (\operatorname{Im} A)^\perp$.

b) On se ramène de même au cas où $0 \in W(A)$ et $|\arg(Av, v)| \leq \theta < \frac{\pi}{2}$, pour tout $v \in \mathbb{C}^d$. Il existe u tel que $(Au, u) = 0$; avec les notations précédentes, on a $(Bu, u) = \operatorname{Re}(Au, u) = 0$, or $(Bv, v) \geq 0$ pour tout v , donc $Bu = \frac{1}{2}(A + A^*)u = 0$. Pour $\varepsilon > 0$ assez petit on a encore $|\arg(e^{i\varepsilon}Av, v)| \leq \frac{\pi}{2}$, donc de même $\frac{1}{2}(e^{i\varepsilon}A + e^{-i\varepsilon}A^*)u = 0$. Par suite $Au = A^*u = 0$. \square

Équation tangentielle de la frontière $\partial W(A)$ de l'image numérique.

Soit $\sigma_0 = x_0 + iy_0 = g^*Ag$, avec $\|g\| = 1$, un point de la frontière. Il existe un demi-plan $\Pi = \{(x, y); ux + vy + w \geq 0\}$ avec $\sigma_0 \in \partial\Pi$ (donc $ux_0 + vy_0 + w = 0$) qui contient l'image numérique : $\Pi \supset W(A)$. On pose

$$G(f) := u \operatorname{Re} f^*Af + v \operatorname{Im} f^*Af + w f^*f.$$

On a $G(g) = 0$ et, puisque $f^*Af/f^*f \in \Pi$, $G(f) \geq 0$ quel que soit $f \in \mathbb{C}^d$. On écrit maintenant la matrice A sous la forme $A = B + iC$, avec B et C auto-adjoints. Alors

$$G(g+h) = G(g) + 2 \operatorname{Re} (u h^*Bg, h) + v h^*Cg + w h^*g) + O(\|h\|^2),$$

ce qui implique $\operatorname{Re} h^*(uB+vC+wI)g = 0$, $\forall h \in \mathbb{C}^d$, et par suite l'égalité $(uB + vC + wI)g = 0$. Or $g \neq 0$, on en déduit

$$P(u, v, w) := \det(uB + vC + wI) = 0. \quad (3)$$

Ainsi toutes les tangentes (et même toute les droites d'appui) de l'image numérique $W(A)$ vérifient l'équation tangentielle $P(u, v, w) = 0$ (de plus, si $P(u, v, w) = 0$, la droite d'équation $ux + vy + w = 0$ possède au moins un point appartenant à l'image numérique. La frontière de l'image numérique est donc contenue dans la courbe algébrique \mathcal{C} d'équation tangentielle $P(u, v, w) = 0$. Notons que P est un polynôme homogène, à coefficients réels, de degré 3, cette courbe est alors dite de classe 3. Ses foyers sont les valeurs propres de A . En effet $\lambda = x_0 + iy_0$ est une valeur propre de A ssi $P(1, i, -\lambda) = 0$, la droite d'équation $x + iy - \lambda = 0$ est tangente à la courbe algébrique et passe par le point de coordonnées (x_0, y_0) . (Rappelons qu'on appelle foyer d'une courbe algébrique, tout point à coordonnées réelles par lequel on peut mener une tangente isotrope à la courbe).

Lorsque le polynôme P est irréductible, l'équation cartésienne $F(x, y)$ de cette courbe s'obtient en éliminant (u, v, w) des relations (redundantes)

$$\begin{aligned} ux + vy + w = 0, \quad P(u, v, w) = 0, \\ \frac{P'_u(u, v, w)}{x} = \frac{P'_v(u, v, w)}{y} = P'_w(u, v, w). \end{aligned}$$

On peut prendre, par exemple,

$$\begin{aligned} F(x, y) := \operatorname{resultant} (P'_u(1, t, -(x+ty)) - x P'_w(1, t, -(x+ty)), \\ P'_v(1, t, -(x+ty)) - y P'_w(1, t, -(x+ty)), t), \end{aligned}$$

ainsi F est un polynôme de degré 6.

Pour obtenir une représentation graphique de la courbe \mathcal{C} on peut procéder de la manière suivante :

soit $w_1(\alpha) \geq w_2(\alpha) \geq w_3(\alpha)$ les racines de $P(\cos \alpha, \sin \alpha, w) = 0$,

- Pour le paramètre α variant de 0 à 2π , on utilise la représentation paramétrique, $j = 1, 2, 3$,

$$x = x_j(\alpha) := \frac{P'_u(\cos \alpha, \sin \alpha, w_j(\alpha))}{P'_w(\cos \alpha, \sin \alpha, w_j(\alpha))}, \quad y = y_j(\alpha) := \frac{P'_v(\cos \alpha, \sin \alpha, w_j(\alpha))}{P'_w(\cos \alpha, \sin \alpha, w_j(\alpha))}.$$

Notons que $z_3(\alpha) := x_3 + iy_3(\alpha) = z_1(\alpha + \pi)$ et $z_2(\alpha) = z_2(\alpha + \pi)$. Lorsque le paramètre α varie de 0 à 2π , le point $z_1(\alpha)$ (tout comme $z_3(\alpha)$) parcourt la frontière de $W(A)$, tandis que le point $z_2(\alpha)$ parcourt deux fois une composante connexe de \mathcal{C} contenue dans $W(A)$.

4 La constante \mathcal{Q}_3 est atteinte.

Lemme 3. *Il existe des matrices $A \in \mathbb{C}^{3,3}$ telles que $\mathcal{Q}_3 = \psi(A)$. Si $\mathcal{Q}_3 > 2$, alors les valeurs propres de ces matrices sont intérieures à $W(A)$, et donc $W(A)$ n'a pas de point anguleux.*

Démonstration. On regarde uniquement le cas $\mathcal{Q}_3 > 2$, l'autre cas étant déjà résolu. Il résulte des propriétés de l'image numérique que $\psi(\lambda A + \mu I) = \psi(A)$ si $\lambda \neq 0$. Il existe donc une suite de matrices A_n à trace nulle et de norme tendant vers 1 telle que $\lim_{n \rightarrow \infty} \psi(A_n) = \mathcal{Q}_3$. Quitte à en extraire si besoin une sous-suite, on peut supposer que A_n converge vers une matrice A , qui sera donc à trace nulle et de norme 1. On a aussi $W(A_n)$ converge vers $W(A)$ au sens de Hausdorff. On regarde les deux cas possibles :

1er cas : A possède une valeur propre $\lambda \in \partial W(A)$. Quitte à effectuer une similitude unitaire (ce qui ne change pas la valeur de $\psi(A)$) on peut supposer d'après le lemme 2 que

$$\text{soit } A = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & & \\ 0 & A' & \end{pmatrix}, \quad \text{soit } A = \begin{pmatrix} -2\lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Le premier cas correspond à λ valeur propre simple, le second cas à une valeur propre double ; ce sont les seuls cas possibles puisque la trace est nulle et la norme vaut 1. Dans les deux cas on peut trouver (au moins pour n assez grand) une suite de matrices H_n inversibles telles que

$$H_n^{-1} A_n H_n = \begin{pmatrix} \mu_n & 0 & 0 \\ 0 & & \\ 0 & A'_n & \end{pmatrix} \quad \text{et} \quad \lim_{n \rightarrow \infty} H_n = I.$$

Il est clair que $\psi(H_n^{-1} A_n H_n) = \psi(A'_n) \leq \mathcal{Q}_2 = 2$ et $\psi(A_n) \leq \|H_n\| \|H_n^{-1}\| \psi(H_n^{-1} A_n H_n)$. Ce cas n'est donc possible que si $\mathcal{Q}_3 = 2$.

2ème cas ; les valeurs propres de A sont toutes intérieures à l'image numérique. Cela implique en particulier que leur barycentre 0 est intérieur. On déduit de la convergence de $W(A_n)$ vers $W(A)$ qu'il existe une suite ε_n convergeant vers 0 telle que $W(A) \subset W((1+\varepsilon_n)A_n)$.

Il existe une suite de polynômes p_n telles que $|p_n| \leq 1$ dans $W(A_n)$ et $\lim_{n \rightarrow \infty} \|p_n(A_n)\| = \mathcal{Q}_3$. Posons $q_n(z) = p_n(\frac{z}{1+\varepsilon_n})$. On a alors $|q_n| \leq 1$ dans $W(A)$, d'où $\psi(A) \geq \|q_n(A)\|$. En choisissant n assez grand pour que les valeurs propres de A_n soient intérieures à $W(A)$, on a

$$q_n(A) = \frac{1}{2\pi i} \int_{\partial W(A)} p_n(z) (z - \frac{1}{1+\varepsilon_n} A)^{-1} dz,$$

$$p_n(A_n) = \frac{1}{2\pi i} \int_{\partial W(A)} p_n(z) (z - A_n)^{-1} dz.$$

On en déduit

$$\|q_n(A) - p_n(A_n)\| \leq \frac{1}{2\pi} \int_{\partial W(A)} \|(z - \frac{1}{1+\varepsilon_n} A)^{-1} - (z - A_n)^{-1}\| dz \rightarrow 0, \quad \text{et}$$

$$\mathcal{Q}_3 \geq \psi(A) \geq \lim_{n \rightarrow \infty} \|q_n(A)\| = \lim_{n \rightarrow \infty} \|p_n(A_n)\| = \mathcal{Q}_3.$$

On a donc $\psi(A) = \mathcal{Q}_3$. □

Avec (2), on en déduit

$$\mathcal{Q}_3 = \max_{A, \alpha_1, \alpha_2} \{\|g(A)\| ; g(z) = \frac{a_A(z) - \alpha_1}{1 - \bar{\alpha}_1 a_A(z)} \frac{a_A(z) - \alpha_2}{1 - \bar{\alpha}_2 a_A(z)}, \quad \alpha_1, \alpha_2 \in \bar{D}\}, \quad (4)$$

a_A étant une représentation conforme de l'intérieur de $W(A)$ sur le disque unité.

Remarque. Soit A la matrice fournie par le lemme ; elle est unitairement semblable à une matrice triangulaire supérieure T , on a donc aussi $\psi(T) = \mathcal{Q}_3$. On remarque que $W(T) = W(e^{i\theta_1} D^{-1} T D)$ et $\psi(T) = \psi(e^{i\theta_1} D^{-1} T D)$, lorsque $D = \text{diag}(1, e^{i\theta_2}, e^{i\theta_3})$. En choisissant les θ_j de manière appropriée, on peut se ramener au cas où les éléments non diagonaux de T sont ≥ 0 . On peut donc se limiter dans (4) à prendre le maximum sur les matrices triangulaires, à trace nulle et à éléments non diagonaux ≥ 0 . On peut aussi rajouter un terme de normalisation (par exemple la somme des éléments non diagonaux égale à 1).

Une autre possibilité est de prendre le maximum sur les matrices de la forme $A = B + iC$, avec B diagonale réelle à trace nulle, C auto-adjointe à trace nulle et à éléments non diagonaux ≥ 0 et de somme 1.

L'expression (4) permet d'envisager un calcul numérique de \mathcal{Q}_3 via un programme d'optimisation portant sur 10 paramètres réels (2 pour les coefficients de B , 2 pour les coefficients diagonaux de C , 2 pour les coefficients non diagonaux de C , 2 pour les arguments de α_1 et α_2 , 2 pour leurs modules (soumis à la contrainte ≤ 1). Pour cela il faut calculer la fonction coût, ce qui exige d'abord la connaissance de $\partial W(A)$, et ensuite le calcul de a_A .

5 Taxonomie des images numériques des matrices 3×3 .

Étudions maintenant la forme des images numériques possibles :

1^{er} cas. (Cas régulier). *Le polynôme P est irréductible et il n'existe pas de point (u_0, v_0, w_0) , $(u_0, v_0) \neq 0$, tel que $P(u_0, v_0, w_0) = P'_w(u_0, v_0, w_0) = 0$, autrement dit : le polynôme $P(u, v, \cdot)$ n'a jamais de racine double, (rappelons que les racines sont toujours réelles). Sans perte de généralité, on peut supposer que $\text{trace}(A) = 0$ (on s'y ramène aisément par une translation). Le polynôme P est alors de la forme*

$$P(u, v, w) = w^3 - 3p(u, v)w + 2q(u, v),$$

avec p et q polynômes homogènes de degrés respectifs 2 et 3. L'hypothèse des racines simples se traduit par $p^3(u, v) > q^2(u, v)$, pour tout $(u, v) \neq 0$. Cela implique en particulier que $p(u, v) > 0$. Rappelons que les racines du polynôme sont données par les formules

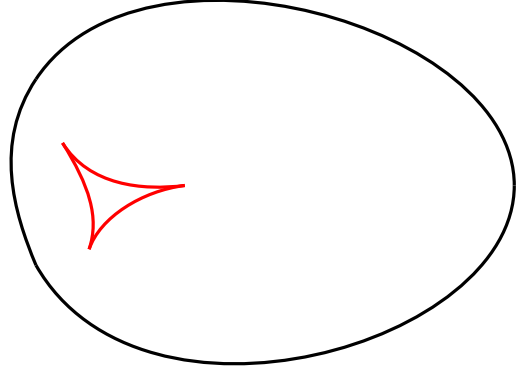
$$\begin{aligned} w_1 &= 2\sqrt{p} \sin\left(\frac{\pi}{3} - \frac{1}{3} \arcsin \frac{q}{p\sqrt{p}}\right), & w_2 &= 2\sqrt{p} \sin\left(\frac{1}{3} \arcsin \frac{q}{p\sqrt{p}}\right), \\ w_3 &= -2\sqrt{p} \sin\left(\frac{\pi}{3} + \frac{1}{3} \arcsin \frac{q}{p\sqrt{p}}\right). \end{aligned}$$

Notons que w_2 est la racine intermédiaire. Posons

$$\begin{aligned} x_j(u, v) &= \frac{-3w_j(u, v)p'_u(u, v) + 2q'_u(u, v)}{3w_j(u, v) - 3p(u, v)} = \frac{P'_u(u, v, w_j)}{P'_w(u, v, w_j)}, \\ y_j(u, v) &= \frac{-3w_j(u, v)p'_v(u, v) + 2q'_v(u, v)}{3w_j(u, v) - 3p(u, v)} = \frac{P'_v(u, v, w_j)}{P'_w(u, v, w_j)}, \end{aligned}$$

et $z_j = x_j + iy_j$. Lorsque α parcourt $[0, 2\pi]$, le point d'affixe $z_1(\cos \alpha, \sin \alpha)$ décrit la frontière $\partial W(A)$, tandis que lorsque α parcourt $[0, \pi]$, le point d'affixe $(z_2(\cos \alpha, \sin \alpha))$ décrit une courbe intérieure \mathcal{C}_i . La courbe algébrique totale $\mathcal{C} = \partial W(A) \cup \mathcal{C}_i$ est de classe 3. Il existe donc 3

tangentes, et 3 seulement, parallèles à toute direction donnée; deux exactement d'entre elles correspondent nécessairement à $\partial W(A)$, puisque $W(A)$ est convexe. Il en résulte que la courbe intérieure a exactement une tangente parallèle à une direction donnée, elle a donc au moins trois points de rebroussement réels, et au plus trois puisque \mathcal{C} est de classe 3 (notons que cette courbe ne peut être réduite à un point $z_0 = x_0 + iy_0$, sinon P serait divisible par $ux_0 + vy_0 + w$ et donc réductible). Pour des raisons algébriques, il existe aussi six autres points de rebroussement qui sont complexes. Les relations de Plücker impliquent alors qu'il n'existe pas d'autres points singuliers et que la courbe algébrique \mathcal{C} est exactement de degré 6. En particulier la frontière de $W(A)$ est analytique.



2^{ème} cas. (Racine double). *Il existe un point (u_0, v_0, w_0) , $(u_0, v_0) \neq 0$, tel que $P(u_0, v_0, w_0) = P'_w(u_0, v_0, w_0) = 0$ et P est irréductible.* Notons que A ne peut avoir des valeurs propres sur la frontière sinon, d'après le lemme 2, le polynôme serait réductible. Nous ne supposons pas ici $\text{trace}(A) = 0$. Quitte à effectuer une rotation, puis une translation, on peut se ramener au cas $(u_0, v_0, w_0) = (1, 0, 0)$; en effectuant ensuite une homothétie, on se ramène au cas où $P(1, 0, w) = w^2(w - 1)$. Puisque l'image numérique ne change pas si on opère une similitude unitaire, on peut supposer que $A = B + iC$ est telle que B est diagonale; on a alors $B = \text{diag}(0, 0, 1)$. Pour la même raison, on peut supposer que la sous-matrice principale 2×2 de C est diagonale, et que les coefficients non diagonaux de C sont positifs. Enfin quitte à remplacer A par $A - c_{33}iI$, on peut supposer que $c_{33} = 0$. On se ramène ainsi au cas où A est de la forme

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + i \begin{pmatrix} a & 0 & \alpha \\ 0 & b & \beta \\ \alpha & \beta & 0 \end{pmatrix}.$$

Notons que $a \neq b$, $\alpha \neq 0$ et $\beta \neq 0$, sinon A serait unitairement semblable à une matrice bloc diagonale, et donc P serait réductible; on peut donc supposer aussi $a > b$. Le polynôme P est alors

$$P(u, v, w) = w^3 + w^2(u + (a+b)v) + w((a+b)uv + (ab - \alpha^2 - \beta^2)v^2) + abuv^2 - (a\beta^2 + b\alpha^2)v^3.$$

Notons que l'on a $P(1, 0, 0) = P'_u(1, 0, 0) = P'_v(1, 0, 0) = P'_w(1, 0, 0) = 0$. Toute cubique à point double étant unicursale, on obtient une représentation paramétrique des solutions de $P(u, v, w) = 0$, qui est

$$u = -(t^3 + (a+b)t^2 + (ab - \alpha^2 - \beta^2)t - (a\beta^2 + b\alpha^2)), \quad v = (t+a)(t+b), \quad w = tv.$$

On a alors

$$\begin{aligned}
 P'_u(u, v, w) &= w^2 + wv((a+b) + abv^2) = v^3, \\
 P'_w(u, v, w) &= 3w^2 + 2w(u+(a+b)v) + (a+b)uv + (ab - \alpha^2 - \beta^2)v^2 \\
 &= v^3 + v((2t^2 - (\alpha^2 + \beta^2)v + (2t+a+b)u) \\
 &= 3w^2 + 2w(u+(a+b)v) + (a+b)uv + (ab - \alpha^2 - \beta^2)v^2 \\
 &= v(v^2 + \alpha^2(t+a)^2 + \beta^2(t+b)^2).
 \end{aligned}$$

On en déduit la représentation paramétrique de la courbe \mathcal{C}

$$\begin{aligned}
 x(t) &= \frac{(t+a)^2(t+b)^2}{(t+a)^2(t+b)^2 + \alpha^2(t+a)^2 + \beta^2(t+b)^2}, \\
 y(t) &= -\frac{u x(t) + w}{v} = -t - \frac{(t+a)(t+b)u(t)}{(t+a)^2(t+b)^2 + \alpha^2(t+a)^2 + \beta^2(t+b)^2}.
 \end{aligned}$$

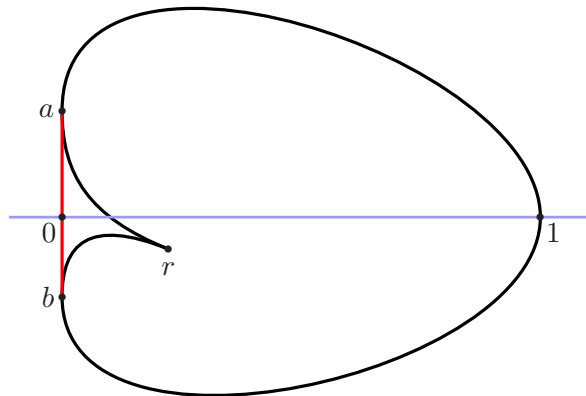
Par dérivation

$$\begin{aligned}
 x'(t) &= \frac{2(t+a)(t+b)(\alpha^2(t+a)^2 + \beta^2(t+b)^2)}{[(t+a)^2(t+b)^2 + \alpha^2(t+a)^2 + \beta^2(t+b)^2]^2}, \\
 y'(t) &= -\frac{u x'(t)}{v} = -\frac{2u(t)(\alpha^2(t+a)^2 + \beta^2(t+b)^2)}{[(t+a)^2(t+b)^2 + \alpha^2(t+a)^2 + \beta^2(t+b)^2]^2}.
 \end{aligned}$$

On remarque qu'il existe un unique r pour lequel $\alpha^2(r+a)^2 + \beta^2(r+b)^2 = 0$, de plus cette racine est simple et $r \in (-a, -b)$; le point $(x(r), y(r))$ est donc un point de rebroussement de première espèce. De plus $u(t) = P(0, 1, t)$ possède trois racines réelles simples $\tau_1 < \tau_2 < \tau_3$; on a $u(-a) = \alpha^2(b-a)$ et $u(-b) = \beta^2(a-b)$ et $u(t) \equiv -t^3$ à l'infini. On en déduit $\tau_1 < -a < \tau_2 < -b < \tau_3$. Rien ne nous permet de situer τ_2 par rapport à r . Le choix $r < \tau_2$ fait dans le tableau de variations ci-dessous est donc arbitraire.

	$-\infty$	τ_1	$-a$	r	τ_2	$-b$	τ_3	∞					
$x'(t)$		-	0	+	0	-	0	+					
$x(t)$	1	\searrow	0	\nearrow	μ	\searrow	0	\nearrow	1				
$y(t)$	0	\nearrow	M	\searrow	a	\searrow	ν	\nearrow	b	\searrow	m	\nearrow	0
$y'(t)$		+	0	-	0	+	0	-	0	+			

La courbe est une cardioïde.
Son enveloppe convexe est $W(A)$.



Dans les cas restants,

soit $P(u, v, w) = (ux_0 + vy_0 + w)q(u, v, w)$ avec q polynôme irréductible de degré 2, alors $W(A)$ est l'enveloppe convexe du point $z_0 = x_0 + iy_0$ et de l'ellipse d'équation tangentielle $q(u, v, w) = 0$,

soit $P(u, v, w) = (u x_0 + v y_0 + w)(u x_1 + v y_1 + w)(u x_2 + v y_2 + w)$, les points $z_j = x_j + i y_j$, $j = 0, 1, 2$ sont alors les valeurs propres de A , $W(A)$ est le triangle de sommets z_0, z_1, z_2 et la matrice A est normale.

6 Calcul de la représentation conforme.

Soit Ω l'intérieur de l'image numérique $W(A)$ d'une matrice correspondant à l'un des deux cas du paragraphe précédent. Quitte à effectuer une translation, on peut supposer A à trace nulle de sorte que $0 \in \Omega$. Soit a une représentation conforme de Ω sur le disque unité telle que $a(0) = 0$. Elle est de la forme $a(z) = z \exp(u + iv)$, avec u, v harmoniques à valeurs réelles. On a alors

$$\Delta u = 0, \quad \text{dans } \Omega, \quad \text{et} \quad u = -\log |z|, \quad \text{sur } \partial\Omega.$$

On en déduit qu'il existe une fonction q définie sur $\partial\Omega$ telle que, pour tout $z \in \Omega$,

$$(u + iv)(z) = \int_{\partial\Omega} q(y) \mathcal{L} \log(y - z) d\sigma(y) + i c \quad \text{et} \quad \int_{\partial\Omega} q(y) d\sigma(y) = 0.$$

On peut faire le choix $c = 0$. La fonction q est à valeurs réelles et s'obtient en résolvant, pour tout $z \in \partial\Omega$,

$$-\log |z| = \int_{\partial\Omega} q(y) \log |y - z| d\sigma(y) \quad \text{et} \quad \int_{\partial\Omega} q(y) d\sigma(y) = 0.$$

7 Matrices dont l'image numérique est le cercle unité.

Lemme 4. *Soit $A \in \mathbb{C}^{3,3}$. Une condition nécessaire et suffisante pour que $W(A)$ soit le disque unité fermé \overline{D} est que A soit unitairement semblable à une matrice de la forme*

$$e^{i\psi} \begin{pmatrix} 0 & 2a & 2b \\ 0 & 0 & 2c \\ 0 & 0 & \zeta \end{pmatrix}, \quad (5)$$

avec $a = \cos \theta$, $b = \sin \theta \cos \varphi$, $c = \sin \theta \sin \varphi$, $\zeta = -a \sin 2\varphi$, $\psi \in \mathbb{R}$ et $\theta, \varphi \in [0, \frac{\pi}{2}]$.

Démonstration. La condition $W(A) = \overline{D}$ implique que, pour tout $u, v \in \mathbb{R}$ on a.

$$P(u, v, \sqrt{u^2 + v^2}) = 0 \quad \text{et} \quad P(u, v, -\sqrt{u^2 + v^2}) = 0.$$

Cette propriété est donc aussi vraie pour tout $u, v \in \mathbb{C}$, ce qui montre que 0 est une racine double de $P(1, i, w) = 0$. Or on a $P(1, i, w) = \det(A - wI)$, donc 0 est valeur propre double de A ; il en résulte que A est unitairement semblable à une matrice de la forme (5), on peut imposer de plus $a \geq 0$, $b \geq 0$, $c \geq 0$, $\psi \in \mathbb{R}$; a priori $\zeta = x_1 + i y_1 \in \mathbb{C}$. On a alors

$$\begin{aligned} P(u, v, w) &= \det \begin{pmatrix} w & a(u-iv) & b(u-iv) \\ a(u+iv) & w & c(u-iv) \\ b(u+iv) & c(u+iv) & w+ux_1+vy_1 \end{pmatrix}, \\ &= w^3 + (ux_1+vy_1)w^2 - (a^2+b^2+c^2)(u^2+v^2)w \\ &\quad + 2abcu(u^2+v^2) - a^2(ux_1+vy_1)(u^2+v^2) \\ &= (w^2 - (u^2+v^2))(w + (ux_1+vy_1)) + (1 - (a^2+b^2+c^2))(u^2+v^2)w \\ &\quad + 2abcu(u^2+v^2) + (1 - a^2)(ux_1+vy_1)(u^2+v^2). \end{aligned}$$

En écrivant que

$$P(u, v, \sqrt{u^2+v^2}) = 0 \quad \text{et} \quad P(u, v, -\sqrt{u^2+v^2}) = 0, \quad \forall u, v \in \mathbb{R},$$

on en déduit $a^2+b^2+c^2 = 1$, $2abc+(1-a^2)x_1 = 0$ et $(1-a^2)y_1 = 0$, le lemme s'ensuit aisément. \square

On recherche maintenant, parmi les matrices ayant pour image numérique le disque unité, lesquelles vérifient $\psi(A) = 2$. D'après le lemme précédent, il suffit de regarder les matrices de la forme (5) avec $\psi = 0$.

Premier cas : $a = 1$ ou $a = 0$ ou $\varphi = 0$. Dans le cas $a = 1$, on est ramené à regarder la matrice

$$A = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \zeta \end{pmatrix}, \quad \text{avec } |\zeta| \leq 1.$$

Il est clair que $\psi(A) = 2$, puisque $\|A\| = 2$. On est dans le même cas lorsque $a = 0$ ou lorsque $\varphi = 0$ puisque les matrices correspondantes sont unitairement semblables à la matrice que l'on vient d'écrire, avec $\zeta = 0$.

On suppose dans tout ce qui suit $a \neq 1$, $a \neq 0$ et $\varphi \neq 0$; on pose

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{ab}{x} & \frac{1-a^2}{x} \\ 0 & \frac{1-a^2}{x} & -\frac{ab}{x} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & y \end{pmatrix}, \quad \text{où} \quad \begin{aligned} x &= \sqrt{(1-a^2)^2 + a^2b^2} = \sin \theta \sqrt{1 - \cos^2 \theta \sin^2 \varphi}, \\ y &= \max(1, \frac{2x}{\sqrt{1-a^2}}) = \max(1, \sqrt{1 - \cos^2 \theta \sin^2 \varphi}), \end{aligned}$$

(notons que H est ainsi le produit d'une matrice élémentaire de Householder par une matrice diagonale). On remarque alors que

$$A = H B H^{-1}, \quad \text{avec } A = \begin{pmatrix} 0 & 2a & 2b \\ 0 & 0 & 2c \\ 0 & 0 & \zeta \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & \frac{b}{x} & \frac{ac^2y}{x} \\ 0 & 0 & 0 \\ 0 & \frac{2c}{y} & \zeta \end{pmatrix},$$

de plus $\|H\| = 2$ et $\|H^{-1}\| = 1$. (Cette écriture a été obtenue en s'inspirant de la démonstration de Badea de $C(D) = 2$ basée sur une décomposition d'Ando (cf. [1]). Nous regardons les cas restants

Deuxième cas : $y \neq \frac{2x}{\sqrt{1-a^2}}$, $a \neq 1$, $a \neq 0$ et $\varphi \neq 0$. Notons que cela implique $y = 1$ et $\cos \theta \sin \varphi > \frac{\sqrt{3}}{2}$. En effet $4x^2 < 1-a^2$ s'écrit aussi $4 \sin^2 \theta + 4 \cos^2 \theta \cos^2 \varphi < 1$, ou encore $\cos \theta \sin \varphi > \frac{\sqrt{3}}{2}$. On a alors

$$B B^* = \begin{pmatrix} 1 - \sin^2 \theta \sin^2 \varphi & 0 & 2 \sin \theta \sin \varphi \cos \varphi \sqrt{1 - \cos^2 \theta \sin^2 \varphi} \\ 0 & 0 & 0 \\ 2 \sin \theta \sin \varphi \cos \varphi \sqrt{1 - \cos^2 \theta \sin^2 \varphi} & 0 & 4 \sin^2 \varphi (1 - \cos^2 \theta \sin^2 \varphi) \end{pmatrix}.$$

On en déduit aussi $\det(I - B B^*) = \sin^2 \theta \sin^2 \varphi (4 \cos^2 \theta \sin^2 \varphi - 3)$. Il en résulte que $I - B B^*$ est définie positive, car ses trois mineurs fondamentaux sont strictement positifs (rappelons que $a \neq 1$ donc $\theta \neq 0$). Il en résulte que $\|B\| < 1$. Par suite on a toujours $\psi(A) < 2$ dans ce deuxième cas.

En effet, si on avait $\psi(A) = 2$, il existerait un produit de Blaschke à au plus deux termes f tel que $\|f(A)\| = 2$; on aurait alors nécessairement $\|f(B)\| = 1$ d'où, puisque $\|B\| < 1$, $f = \text{constante}$, ce qui contredit $\|f(A)\| = 2$.

Troisième cas : $1 < y = \frac{2x}{\sqrt{1-a^2}} < 2$, $a \neq 1$, $a \neq 0$ et $\varphi \neq 0$. Cela s'écrit de manière équivalente $0 < \cos \theta \sin \varphi < \frac{\sqrt{3}}{2}$ et $\theta \neq 0$.

On a alors $b \neq 0$ et, par un calcul simple,

$$B^*B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{z^2 x^2}{b^2} \end{pmatrix}.$$

Notons que

$$\begin{aligned} b^2 - x^2 z^2 &= \sin^2 \theta \cos^2 \varphi (1 - 4(\sin^2 \theta + \cos^2 \theta \cos^2 \varphi) \cos^2 \theta \sin^2 \varphi) \\ &= \sin^2 \theta \cos^2 \varphi (1 - 2 \cos^2 \theta \sin^2 \varphi)^2 \geq 0. \end{aligned}$$

On en déduit $\|B\| = 1$. Supposons maintenant que l'on a $\psi(A) = 2$. Il existe alors un produit de Blaschke $f(z) = \frac{z-\lambda}{1-\lambda z} \frac{z-\mu}{1-\mu z}$, avec $|\mu| \leq |\lambda| \leq 1$, tel que $\|f(A)\| = 2$. On peut donc trouver deux vecteurs u et $v \in \mathbb{C}^3$ avec $\|u\| = \|v\| = 1$ tels que $u^* f(A) v = 2$. On a alors $2 = u^* H f(B) H^{-1} v$, avec $\|u^* H\| \leq 2$, $\|f(B)\| \leq 1$ et $\|H^{-1} v\| \leq 1$; il en résulte que $\|u^* H\| = 2$, $\|f(B)\| = 1$ et $\|H^{-1} v\| = 1$. Compte-tenu de la forme de H , cela implique que $u^* H = 2\rho(1, 0, 0) = 2\rho e_1^*$ et $H^{-1} v = \sigma e_2$, avec $|\rho| = |\sigma| = 1$. En utilisant que $\|f(B)\| = 1$, cela montre finalement que l'on a $f(B) e_2 = \nu e_1$, avec $|\nu| = 1$. On utilise maintenant la représentation de Newton

$$f(B) = f_0 I + f_1 B + f_2 B^2, \quad \text{avec } f_0 = f(0), \quad f_1 = f'(0), \quad f_2 = f[0, 0, \zeta],$$

on obtient aisément

$$\begin{aligned} 0 &= e_2^* f(B) e_2 = f_0, & 0 &= e_3^* f(B) e_2 = \frac{2c}{y} (f_1 + \zeta f_2), \\ \nu &= e_1^* f(B) e_2 = \frac{1}{x} (b f_1 + 2 a c^3 f_2). \end{aligned}$$

On en déduit $\mu = 0$, $f_1 = -\lambda$, $f_2 = \frac{1-|\lambda|^2}{1-\lambda\zeta}$; notons que $\theta \neq 0$ et donc $c \neq 0$, on a alors $f_1 + \zeta f_2 = 0$, d'où $\lambda = \zeta$, $f_2 = 1$ et $f_1 = -\zeta$. Par suite

$$\nu = \frac{1}{x} (-b\zeta + 2ac^3) = \frac{2 \cos \theta \sin \varphi (\sin^2 \theta \sin^2 \varphi + \cos^2 \varphi)}{\sqrt{1 - \cos^2 \theta \sin^2 \varphi}} = 2 \cos \theta \sin \varphi \sqrt{1 - \cos^2 \theta \sin^2 \varphi}.$$

Cela n'est compatible avec $|\nu| = 1$ que lorsque $\cos \theta \sin \varphi = \frac{1}{\sqrt{2}}$; dans ce dernier cas on a effectivement $u^* f(A) v = 2\nu$ en prenant $u = e_1$ et $v = H e_2$. Donc dans ce troisième cas on a $\psi(A) = 2$ si, et seulement si, $\cos \theta \sin \varphi = \frac{1}{\sqrt{2}}$.

Quatrième cas : $y = \frac{2x}{\sqrt{1-a^2}} = 1$, $a \neq 1$, $a \neq 0$ et $\varphi \neq 0$. Ce cas correspond à $\cos \theta \sin \varphi = \frac{\sqrt{3}}{2}$ et $\theta \neq 0$; on a donc $\theta \in]0, \frac{\pi}{6}]$, $\varphi \in]\frac{\pi}{3}, \frac{\pi}{2}]$, $x = \frac{\sin \theta}{2}$, $y = 1$. Comme on a $2 \cos \varphi < 1$, on peut trouver $t \in]0, \pi/2]$ tel que $\cos t = 2 \cos \varphi$. On a alors

$$4c^2 + 4 \cos^2 \varphi = 4(\sin^2 \theta \sin^2 \varphi + \cos^2 \varphi) = 1, \quad \text{donc } \frac{2c}{y} = \sin t,$$

$$\zeta = -2 \cos \theta \sin \varphi \cos \varphi = -\frac{\sqrt{3}}{2} \cos t, \quad \frac{ac}{x} = \sqrt{3}, \quad \text{donc } \frac{ac^2 y}{x} = \frac{\sqrt{3}}{2} \sin t.$$

D'où

$$B = \begin{pmatrix} 0 & \cos t & \frac{\sqrt{3}}{2} \sin t \\ 0 & 0 & 0 \\ 0 & \sin t & -\frac{\sqrt{3}}{2} \cos t \end{pmatrix} \quad \text{et} \quad B^*B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3/4 \end{pmatrix}.$$

On en déduit comme précédemment $\|B\| = 1$ et, en supposant $\psi(A) = 2$, il existe un produit de Blaschke à deux termes f tel que $\|f(A)\| = 2$ et deux vecteurs u et $v \in \mathbb{C}^3$ avec $\|u\| = \|v\| = 1$ tels que $\|u^*H\| = 2$, $\|f(B)\| = 1$ et $\|H^{-1}v\| = 1$. Compte-tenu de la forme de H , cela implique comme précédemment que $u^*H = 2\rho e_1^*$ avec $|\rho| = 1$, mais maintenant on sait seulement que $w := H^{-1}v = (0, \alpha, \beta)^T$, avec $|\alpha|^2 + |\beta|^2 = 1$. En utilisant que $\|f(B)\| = 1$, cela montre finalement que l'on a $f(B)w = \nu e_1$, avec $|\nu| = 1$.

En reprenant la représentation de Newton utilisée au cas précédent on obtient

$$\begin{aligned} 0 &= e_2^* f(B)w = f_0 \alpha, & 0 &= e_3^* f(B)w = (f_1 - \frac{\sqrt{3}}{2} \cos t f_2)(\alpha \sin t - \beta \frac{\sqrt{3}}{2} \cos t) + f_0 \beta, \\ \nu &= e_1^* f(B)w = f_1(\alpha \cos t + \beta \frac{\sqrt{3}}{2} \sin t) + f_2 \frac{\sqrt{3}}{2} \sin t(\alpha \sin t - \beta \frac{\sqrt{3}}{2} \cos t). \end{aligned}$$

On a trois sous-cas possibles

- $f_0 = 0$ et $f_1 + \zeta f_2 = 0$. Alors $\mu = 0$, $f_1 = -\lambda = -\zeta = \frac{\sqrt{3}}{2} \cos t$, $f_2 = \frac{1-|\lambda|^2}{1-\lambda\zeta} = 1$. Par suite $\nu = \frac{\sqrt{3}}{2} \alpha$, ce qui est incompatible avec $|\nu| = 1$.

- $f_0 = 0$ et $f_1 + \zeta f_2 \neq 0$. Alors $\alpha \sin t - \beta \frac{\sqrt{3}}{2} \cos t = 0$ donc $\alpha = \eta \frac{\sqrt{3}}{2} \cos t$, $\beta = \eta \sin t$, avec $|\eta| = (\frac{3}{4} \cos^2 t + \sin^2 t)^{-1/2}$. Par suite $\nu = \eta \frac{\sqrt{3}}{2} f_1$, or on a $|f_1| \leq 1$ et $\eta \frac{\sqrt{3}}{2} < 1$, ce qui est aussi incompatible avec $|\nu| = 1$.

- $\alpha = 0$. On peut prendre $w = e_3$, i.e. $\beta = 1$. On doit alors avoir $f_0 = \frac{\sqrt{3}}{2} \cos t (f_1 - \frac{\sqrt{3}}{2} \cos t f_2) = -\zeta (f_1 + \zeta f_2) = f_0 - f(\zeta)$. Par suite $f(\zeta) = 0$, d'où $\mu = \zeta$, $f_0 = \lambda \zeta$, $(f_1 + \zeta f_2) = -f_0/\zeta = -\lambda$. En reportant dans ν on obtient $\nu = \frac{\sqrt{3}}{2} \sin t (f_1 + \zeta f_2) = -\lambda \frac{\sqrt{3}}{2} \sin t$ ce qui est encore incompatible avec $|\nu| = 1$.

Dans ce dernier cas on n'a donc jamais $\psi(A) = 2$.

Lemme 5. Parmi les matrices 3×3 , les matrices telles que $\psi(A) = 2$ et $W(A) = \overline{D}$ sont les matrices qui sont unitairement semblables à une matrice de l'une des deux familles suivantes

$$e^{i\psi} \begin{pmatrix} 0 & \frac{2}{\sqrt{2-t^2}} & \frac{t(\sqrt{2-2t^2})}{\sqrt{2-t^2}} \\ 0 & 0 & \sqrt{2-2t^2} \\ 0 & 0 & -t \end{pmatrix}, \quad \text{avec } t \in [0, 1], \psi \in \mathbb{R},$$

$$\begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \zeta \end{pmatrix}, \quad \text{avec } \sigma \in \mathbb{C}, |\zeta| \leq 1.$$

8 Annexe. Un problème de Pick-Nevanlinna.

Avant d'aborder ce problème, rappelons que tous les automorphismes holomorphes du disque unité (ouvert) D du plan complexe sont de la forme $z \mapsto m(z)$ (transformation de Möbius)

$$m(z) = e^{i\theta} \frac{z - \alpha}{1 - \bar{\alpha}z}, \quad \alpha \in D, \theta \in \mathbb{R}.$$

Par ailleurs, on appelle produit de Blaschke à k termes les fonctions de la forme

$$b(z) = e^{i\theta} \prod_{j=1}^k \frac{z - \alpha_j}{1 - \bar{\alpha}_j z}, \quad \alpha_j \in D, \theta \in \mathbb{R}.$$

Les transformations de Möbius sont donc des cas particuliers des produits de Blaschke. De plus on vérifie assez aisément que si m est une transformation de Möbius, et si b est un produit de

Blaschke à k termes, alors $m \circ b$ et $b \circ m$ sont des produits de Blaschke à k termes. Remarquons la propriété $|b(z)| = 1$, pour tout $z \in \partial D$. Enfin on note $H^\infty(D)$ l'espace des fonctions holomorphes et bornées dans D .

Le problème de Pick-Nevanlinna est le suivant :

On se donne $\zeta_0, \zeta_1, \dots, \zeta_k$, $k+1$ points distincts de D ,
et f_0, f_1, \dots, f_k , $k+1$ nombres complexes.

Trouver $f \in H^\infty(D)$ qui réalise le minimum de $\|f\|_{L^\infty(D)}$
sous les contraintes $f(\zeta_j) = f_j$, $j = 0, \dots, k$. (6)

Théorème 6. *Le problème de Pick-Nevanlinna admet une solution et une seule. De plus cette solution est de la forme $f = Cb$, où C est la valeur du minimum et b un produit de Blaschke à, au plus, k termes.*

Démonstration. a) Existence. L'ensemble des fonctions holomorphes réalisant les contraintes n'est pas vide; il y a au moins le polynôme d'interpolation de Lagrange. Soit M sa norme; l'existence se déduit alors de la compacité (dans $C^m(K)$ quel que soit K compact avec $K \subset D$) des fonctions $f \in H^\infty(D)$ telles que $\|f\|_{L^\infty(D)} \leq M$.

b) Forme et unicité. Le cas $C = 0$ est trivial; dans le cas contraire, quitte à diviser les f_j par C , on peut supposer que f réalise le minimum et que $C = 1$. On procède alors par récurrence sur k . Pour $k = 0$, il est clair que l'unique solution est $f(z) = f_0$, fonction constante. Supposons le résultat montré jusqu'à $k - 1$ et $C = 1$. Dans le cas trivial où tous les f_j sont égaux, la solution est encore $f(z) = f_0$, fonction constante. Autrement, par le principe du maximum, on a nécessairement $|f_k| < 1$. On pose alors

$$\tilde{f}(z) = (m \circ f)(z), \quad \tilde{f}_j = m(f_j), \quad \text{où l'on a posé } m(z) = \frac{z - z_k}{1 - \bar{z}_k z}.$$

On a clairement $\tilde{f} \in H^\infty(D)$, $\tilde{f}(\zeta_j) = \tilde{f}_j$ et

$$1 = \|\tilde{f}\|_{L^\infty(D)} = \min\{\|g\|_{L^\infty(D)}; g \in H^\infty(D), g(\zeta_j) = \tilde{f}_j, j = 0, \dots, k\}.$$

En effet, s'il existait $g \in H^\infty(D)$, avec $\|g\|_{L^\infty(D)} < 1$ et $g(\zeta_j) = \tilde{f}_j$, $j = 0, \dots, k$, on aurait $m^{-1} \circ g \in H^\infty(D)$, $\|m^{-1} \circ g\|_{L^\infty(D)} < 1$, $(m^{-1} \circ g)(\zeta_j) = f_j$, $j = 0, \dots, k$, ce qui serait contraire à l'hypothèse 1 valeur minimale.

Notons que $\tilde{f}(\zeta_k) = 0$. On pose maintenant

$$h(z) = \frac{\tilde{f}(z)(1 - \bar{\zeta}_k z)}{z - \zeta_k}, \quad h_j = \frac{\tilde{f}_j(1 - \bar{\zeta}_k \zeta_j)}{\zeta_j - \zeta_k}, \quad j = 0, \dots, k-1.$$

On a $h \in H^\infty(D)$, $h(\zeta_j) = h_j$, $j = 0, \dots, k-1$, et $\|h\|_{L^\infty(D)} = \|h\|_{L^\infty(\partial D)} = \|\tilde{f}\|_{L^\infty(D)} = 1$. Par un raisonnement similaire à précédemment, on obtient

$$1 = \|h\|_{L^\infty(D)} = \min\{\|g\|_{L^\infty(D)}; g \in H^\infty(D), h(\zeta_j) = h_j, j = 0, \dots, k-1\}.$$

L'hypothèse de récurrence nous donne que h est unique, et est un produit de Blaschke à au plus $k-1$ termes, donc \tilde{f} est unique, et est un produit de Blaschke à au plus k termes, par suite $f = m^{-1} \circ \tilde{f}$ est unique et est un produit de Blaschke à au plus k termes. □

Remarques.

1) Il n'est pas difficile de modifier la démonstration pour voir que le théorème reste valable si on considère le problème généralisé suivant :

On se donne $\zeta_0, \zeta_1, \dots, \zeta_k, k+1$ points, distincts ou non, de D ,
et $f_0, f_1, \dots, f_k, k+1$ nombres complexes.

Trouver $f \in H^\infty(D)$ qui réalise le minimum de $\|f\|_{L^\infty(D)}$
sous les contraintes $f[\zeta_0, \dots, \zeta_j] = f_j, j = 0, \dots, k$.

2) Si $|\alpha| = 1$ alors $\frac{z-\alpha}{1-\bar{\alpha}z} = -\alpha$ est de module 1. En admettant d'accepter des $\alpha_j \in \overline{D}$, on peut donc considérer qu'un produit de Blaschke à au plus k termes est aussi un produit de Blaschke à k termes.

References

- [1] M. Crouzeix, Bounds for analytic functions of matrices, *Int. Equ. Op. Th.* **48**, 461–477, 2004.
- [2] M. Crouzeix, Numerical range and functional calculus in Hilbert space, preprint, available from <http://perso.univ-rennes1.fr/michel.crouzeix/>

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`