

Sur la convergence locale de la méthode des projections alternées

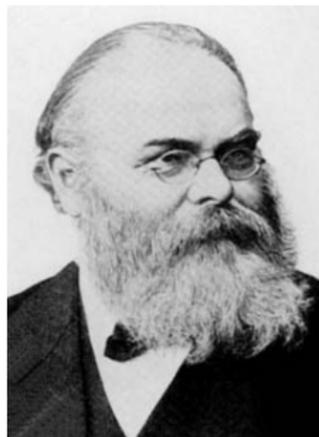
Dominikus Noll



Projections alternées inventées par Hermann Schwarz en 1869

H. Schwarz. Über einen Grenzübergang durch alternirendes Verfahren.

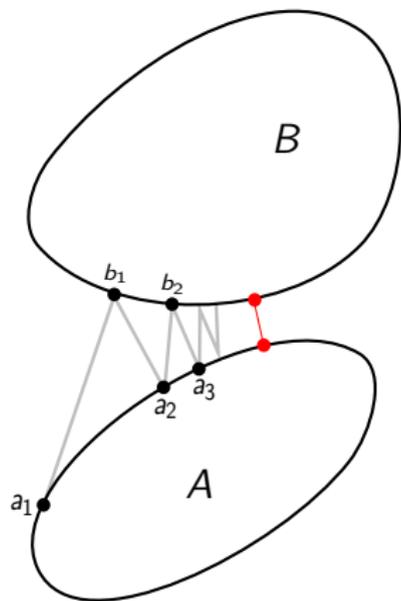
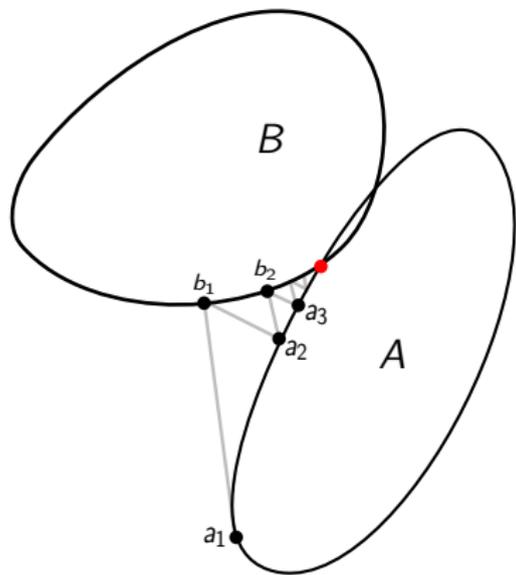
Vierteljahresschrift der Naturforschenden Gesellschaft in Zürich, 15(1870), pp. 272–286.

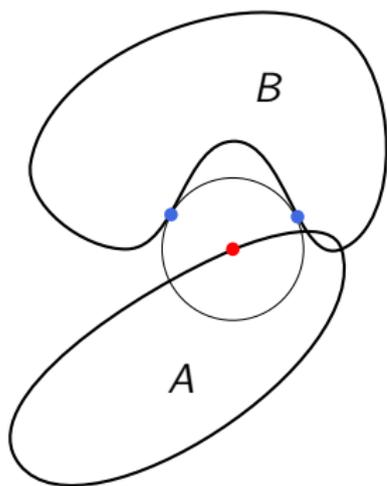
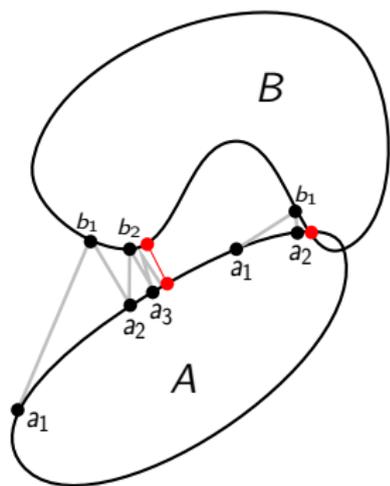


- Proposée pour résoudre rigoureusement le problème de Dirichlet
- Première méthode de décomposition de domaines
- Travaux revisités par P.-L. Lions 1978, 1988-89.

Projections alternées dans le cas convexe

- Schwarz, Sobolev, v. Neumann : sous-spaces affines
- W. Cheney, A. Goldstein : Convergence pour ensembles convexes fermés dans \mathbb{R}^n . 1959
- L.M. Bregman : Convergence faible pour convexes fermés dans l'espace de Hilbert. 1965
- H.H. Bauschke : Cas convexe essentiellement résolu en 1995.
- H. Hundal. Echec de la convergence forte, 2002.





Données : ensembles fermés A, B dans \mathbb{R}^n

$$A \cap B \neq \emptyset$$

Cherche : solution x du problème de faisabilité

$$x \in A \cap B$$

Méthode :

$$b_1 \in P_B(a_1), a_2 \in P_A(b_1), b_2 \in P_B(a_2), a_3 \in P_A(b_2), \dots$$

ou encore

$$a_1 \xrightarrow{P_B} b_1 \xrightarrow{P_A} a_2 \xrightarrow{P_B} b_2 \xrightarrow{P_A} \dots$$

Projections alternées dans le cas non-convexe :

- Existe-t-il des applications ?
- Conditions assurant la convergence locale ?
- Echec de convergence possible ?

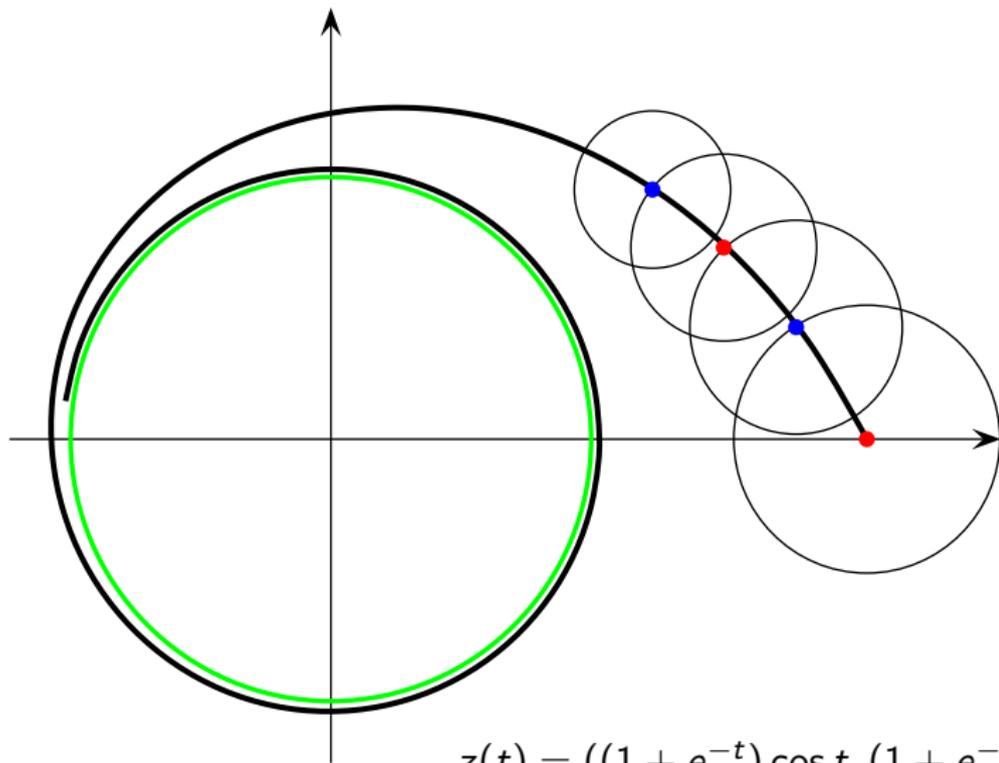
Echec de convergence

Théorème

(Combettes, Trussell 1990). Soient A, B des fermés dans \mathbb{R}^n . Supposons que la suite alternée a_k, b_k est bornée et vérifie $a_k - b_k \rightarrow 0$. Alors l'ensemble des points d'accumulation de a_k, b_k est ou un singleton, ou un compact continu.

Théorème

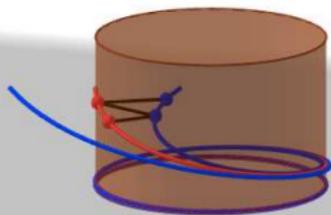
(Bauschke, Noll 2013). Le cas d'un compact continu non-singleton peut se produire.



$A = \text{rouge} \cup \mathbb{S}^1$

$B = \text{bleu} \cup \mathbb{S}^1$

$A \cap B = \mathbb{S}^1$



$$A = \{(\cos t, \sin t, s) : 0 \leq s \leq 1, 0 \leq t \leq 2\pi\}$$

$$B = \{((1 + e^{-t}) \cos t, (1 + e^{-t}) \sin t, e^{-t/2}) : 0 \leq t \leq \infty\}$$

Bauschke, Noll (2014, Archiv der Math.)

Existe-t-il des applications ?

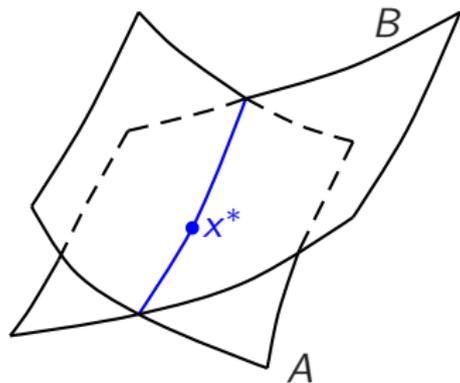
Applications des projections alternées non-convexes :

- Démosaïquage de couleurs
- Débruitage de séries chronologiques (algorithme de Cadzow)
- Tomographie
- Placement de pôles (en contrôle)
- Synthèse de lois de commande en feedback (contrôle)
- Ebauche de tracements de routes (Road profile design)
- Restitution de blocks d'images JPEG et MPEG perdus
- Faisabilité affine creuse (correction de fautes dans des codes linéaires)
- Empilements denses dans des variétés Grassmanniennes (télécommunication)
- Algorithme EM pour les lois normales
- Problème de phase (interférométrie, cristallographie, ...)

Convergence locale

Théorème

(A.S. Lewis, J. Malick 2008). Soient A, B des sous-variétés de classe C^2 de \mathbb{R}^n , dont l'intersection est transversale en $x^* \in A \cap B$. Alors il existe un voisinage U de x^* tel que toute suite alternée a_k, b_k qui rentre dans U converge vers un $a^* \in A \cap B$ avec vitesse R -linéaire.



Transversalité

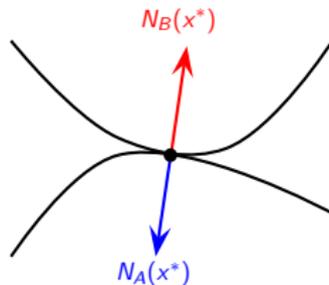
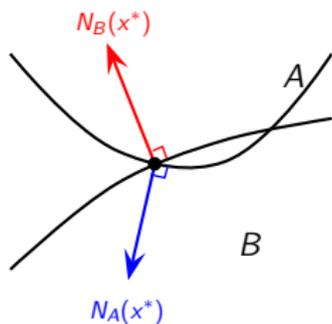
$$T_A(x^*) + T_B(x^*) = \mathbb{R}^n$$

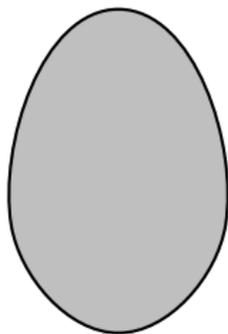
Théorème

(A.S. Lewis, R. Luke, J. Malick 2009). *Supposons le suivant :*

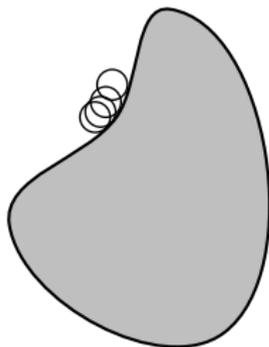
- 1 *Il existe $x^* \in A \cap B$ tel que $N_A(x^*) \cap -N_B(x^*) = \{0\}$ (remplace transversalité);*
- 2 *B est super-régulier (remplace convexité).*

Alors il exist un voisinage U de x^ tel que toute suite alternée a_k, b_k qui rentre dans U converge vers un $a^* \in A \cap B$ avec vitesse R -linéaire.*

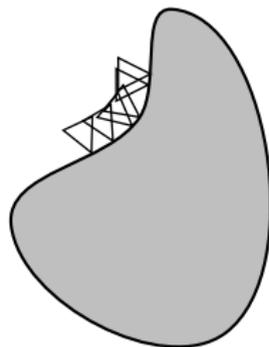




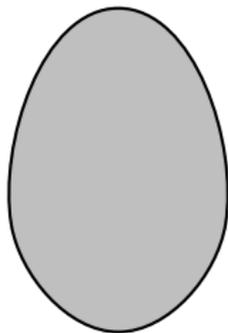
convexe



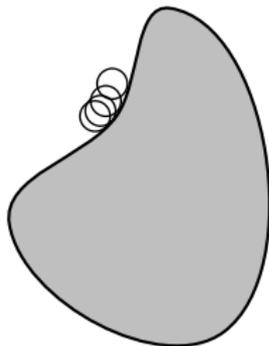
prox-régulier



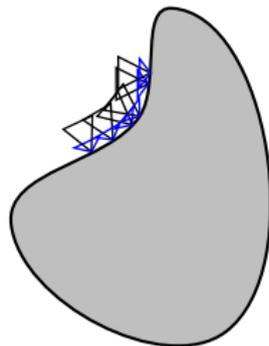
super-régulier



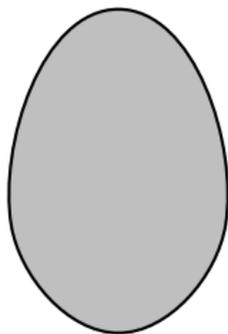
convexe



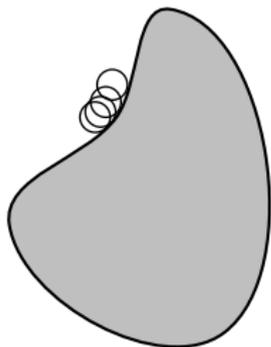
prox-régulier



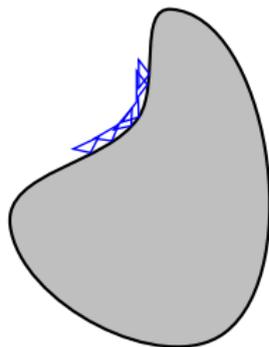
super-régulier



convexe



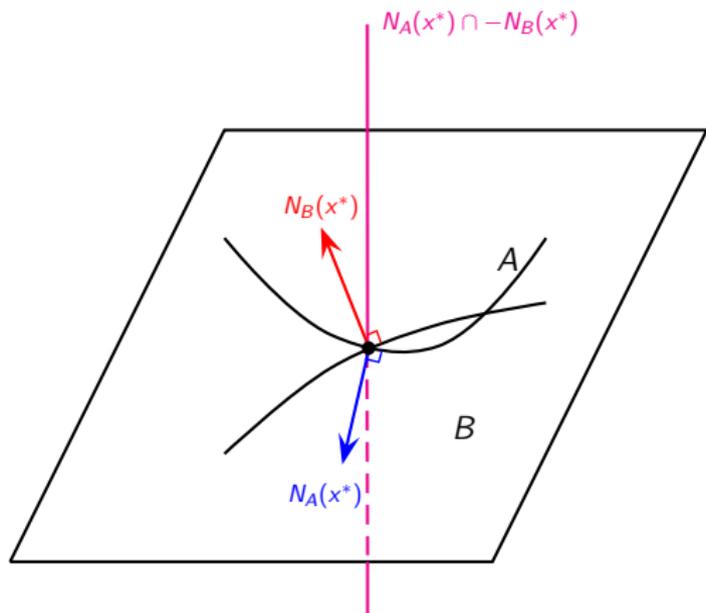
prox-régulier



super-régulier

H.H. Bauschke, D.R. Luke, H.M. Phan, X. Wang (2013). Restricted normal cones and the method of alternating projections.

Set-valued Var. Anal.



Cône normal restreint - la solution ? :

$N_A^B(x^*) =$ vecteur normal à A en x^* qui pointe dans B

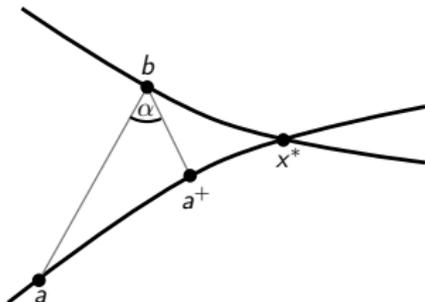
Transversalité en x^* remplacée par :

$$N_A^B(x^*) \cap -N_B^A(x^*) \subset \{0\}$$

Marche mieux, **mais ce n'est pas la fin.**

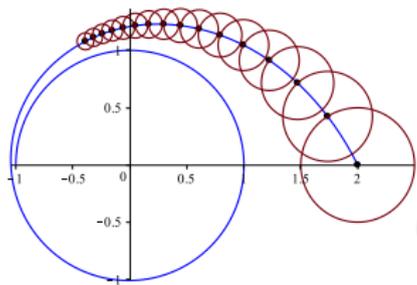
Définition

(Noll, Rondepierre 2013). Transversalité est si α ne se rapproche pas de 0° au voisinage de x^* .

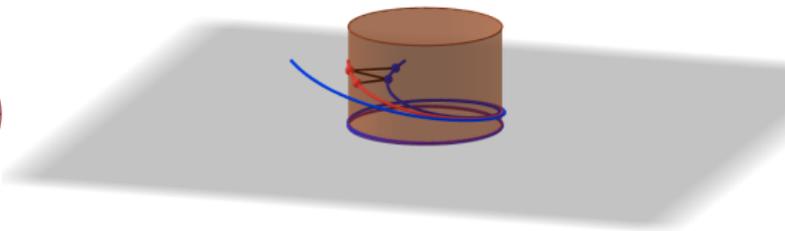


Que se passe-t-il si l'intersection est tangentielle ?

- L'échec de convergence, est-il dû à un manque de régularité ?
- ou parce que l'intersection est (trop) tangentielle ?

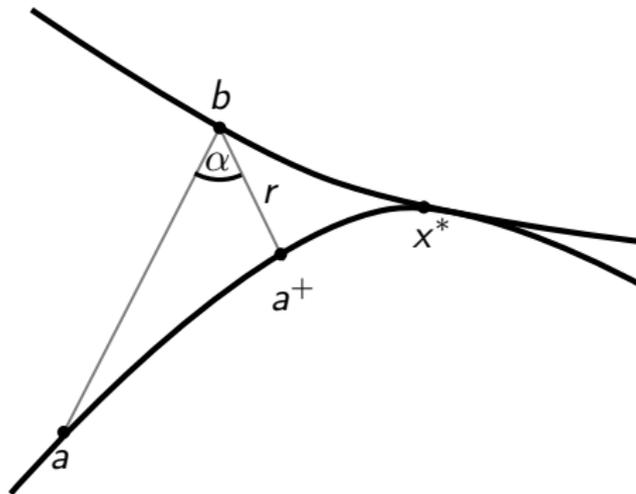


$\alpha \approx 180^\circ$ (transversale)
 Manque de régularité

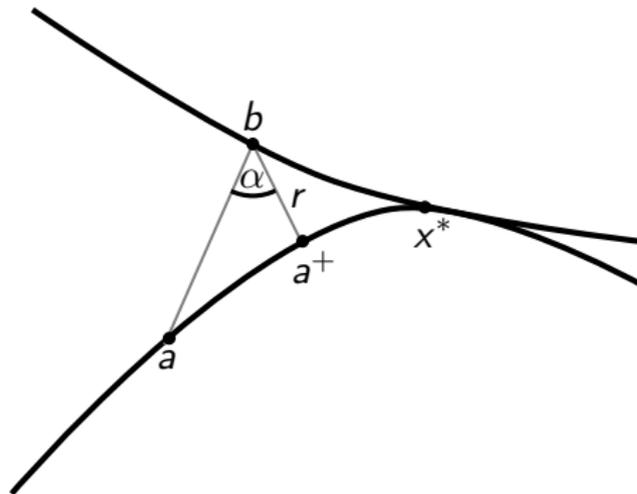


intersection tangentielle
 Régularité bonne

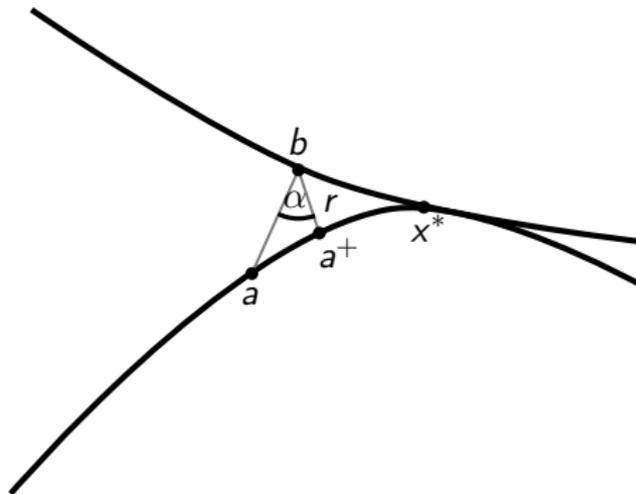
Intersection tangentielle :



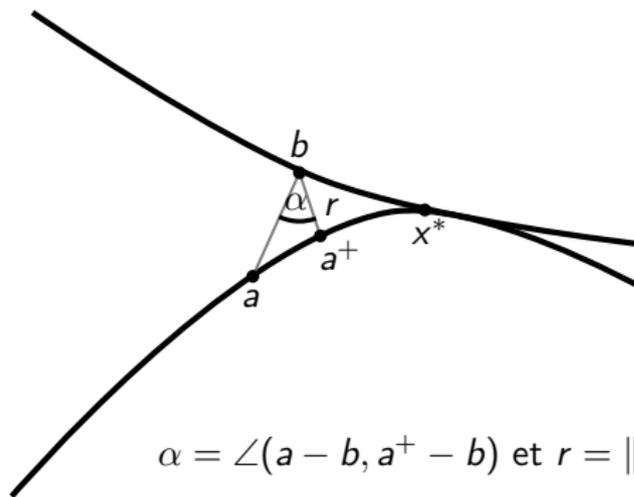
Intersection tangentielle :



Intersection tangentielle :



Intersection tangentielle :



$$\alpha = \angle(a - b, a^+ - b) \text{ et } r = \|b - a^+\|$$

tendent tous les deux vers 0 quand $a, b, a^+ \rightarrow x^*$

Définition

(Noll, Rondepierre 2013). Les fermés A, B vérifient la condition d'angle en $x^* \in A \cap B$ s'il existe $\gamma > 0$, $\omega \in [0, 2)$ et un voisinage U de x^* tel que pour tout block $a \xrightarrow{P_B} b \xrightarrow{P_A} a^+$ dans U on a

$$\frac{1 - \cos \alpha}{r^\omega} \geq \gamma$$

- Intersection tangentielle signifie α et r tendent vers 0.
- Condition d'angle signifie qu'il y a un lien entre les deux. L'angle ne s'annule pas trop vite.
- $\omega = 0$ transversalité ($\alpha \not\rightarrow 0$, $r \rightarrow 0$).

Théorème

(Noll, Rondepierre 2013). *Supposons qu'il existent $x^* \in A \cap B$ et $\omega \in (0, 2)$ tels que*

- 1 *la condition d'angle pour ce ω est vérifiée en $x^* \in A \cap B$.*
- 2 *B est régulier en x^* au sens de Hölder par rapport à A avec $\omega/2$.*

Alors il existe un voisinage U de x^ tel que toute suite alternée a_k, b_k qui rentre dans U converge vers un $a^* \in A \cap B$ avec vitesse*

$$\|a_k - a^*\| = \mathcal{O}\left(k^{-\frac{2-\omega}{2\omega}}\right), \quad \|b_k - a^*\| = \mathcal{O}\left(k^{-\frac{2-\omega}{2\omega}}\right)$$

Cas $\omega = 0$ donne convergence R-linéaire comme avant

Théorème

(Noll, Rondepierre 2013). Soient A, B des fermés sous-analytiques. Alors en tout $x^* \in A \cap B$ la condition d'angle est vérifiée pour un certain $\omega \in [0, 2)$.

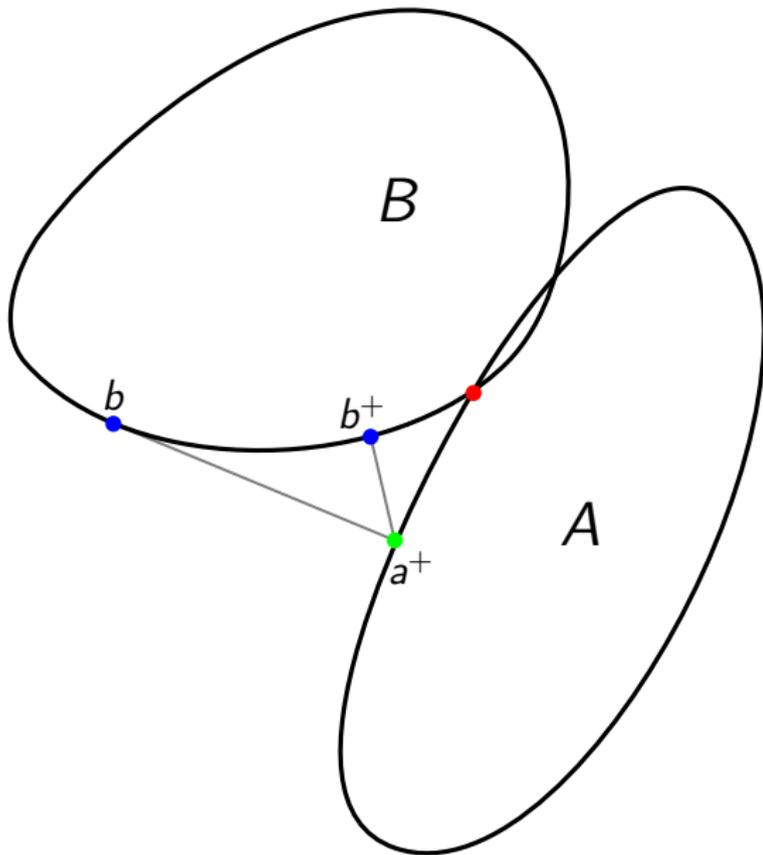
Ensemble semi-analytique :

$$A = \bigcup_{i=1}^N \bigcap_{j=1}^M \{x \in \mathbb{R}^n : \phi_{ij}(x) = 0, \psi_{ij}(x) > 0\}$$

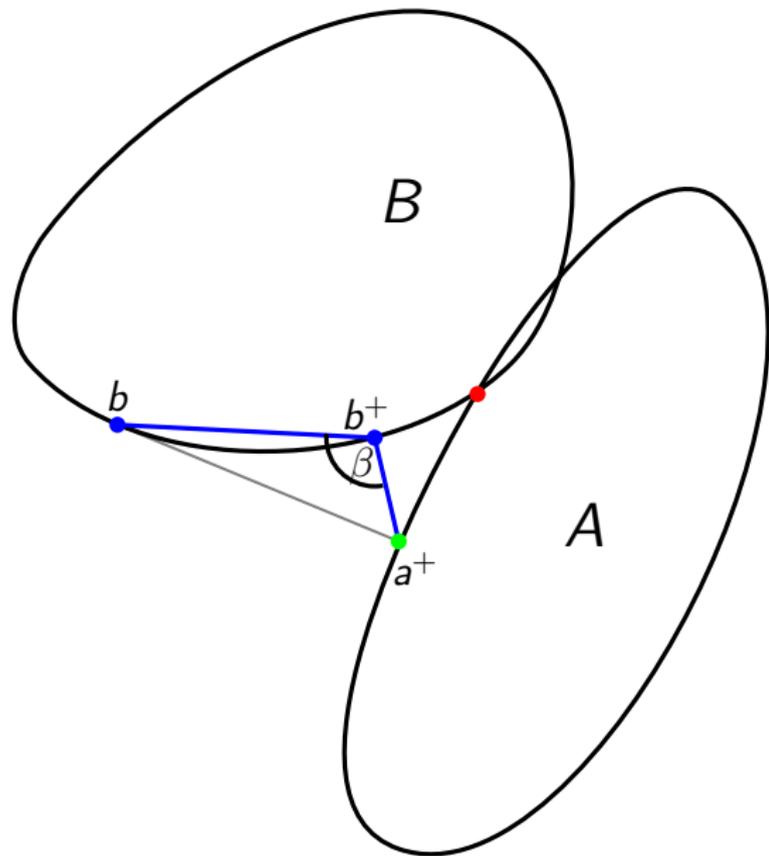
avec ϕ_{ij}, ψ_{ij} réel-analytiques.

A sous-analytique $\iff \forall a \in A \exists r > 0 \exists \mathcal{A}$ semi-analytique borné
 $A \cap B(a, r) = \{x : (x, y) \in \mathcal{A}\}$

Que signifie la Hölder régularité ?

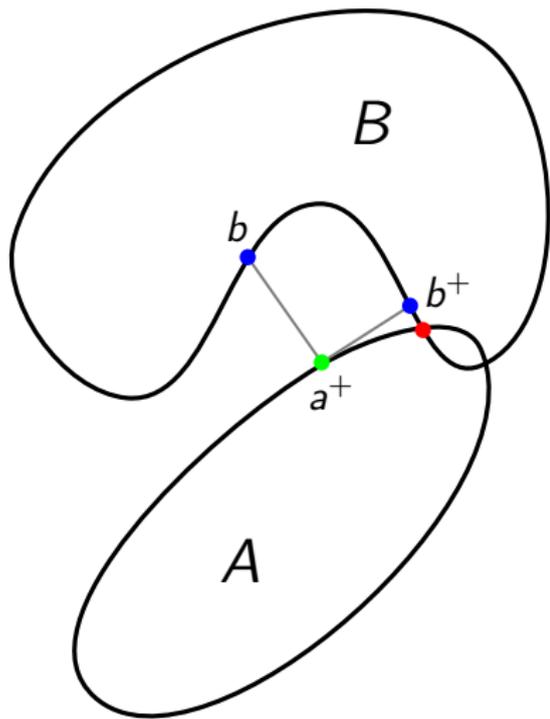


B convexe

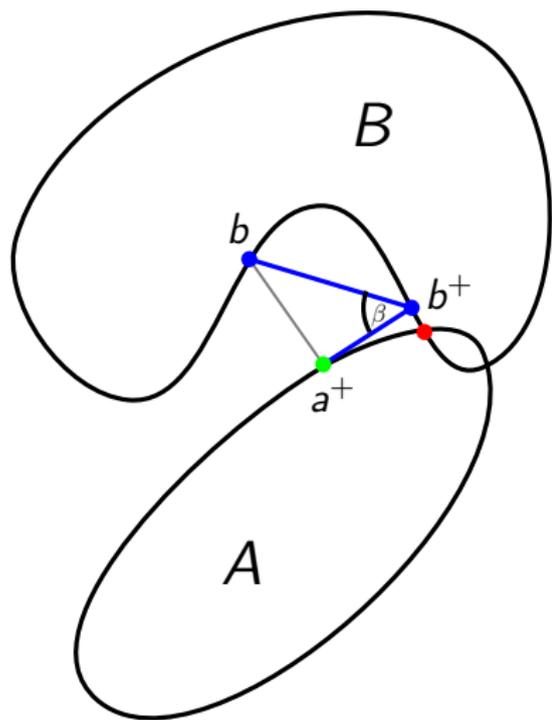


B convexe

$$\beta \geq 90^\circ$$

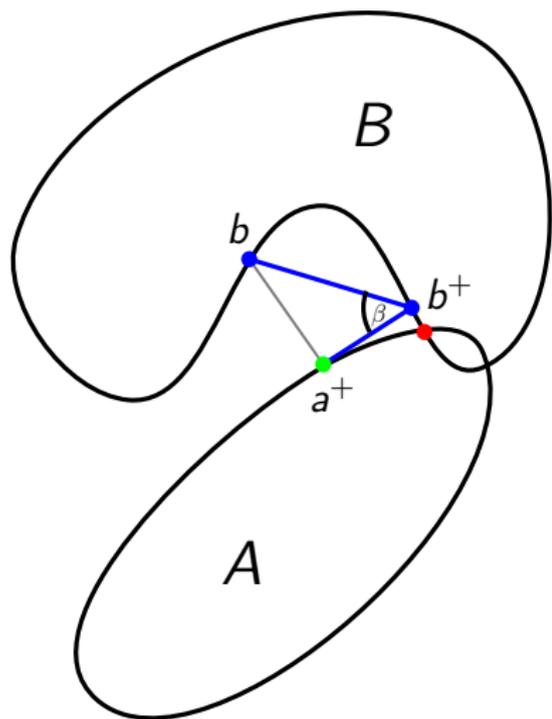


B non-convexe



B non-convexe

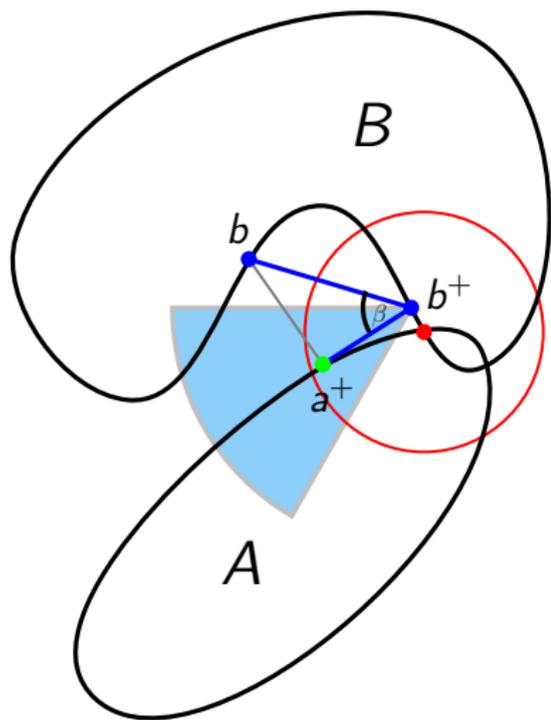
$\beta < 90^\circ$ possible



B non-convexe

si B super-régulier :

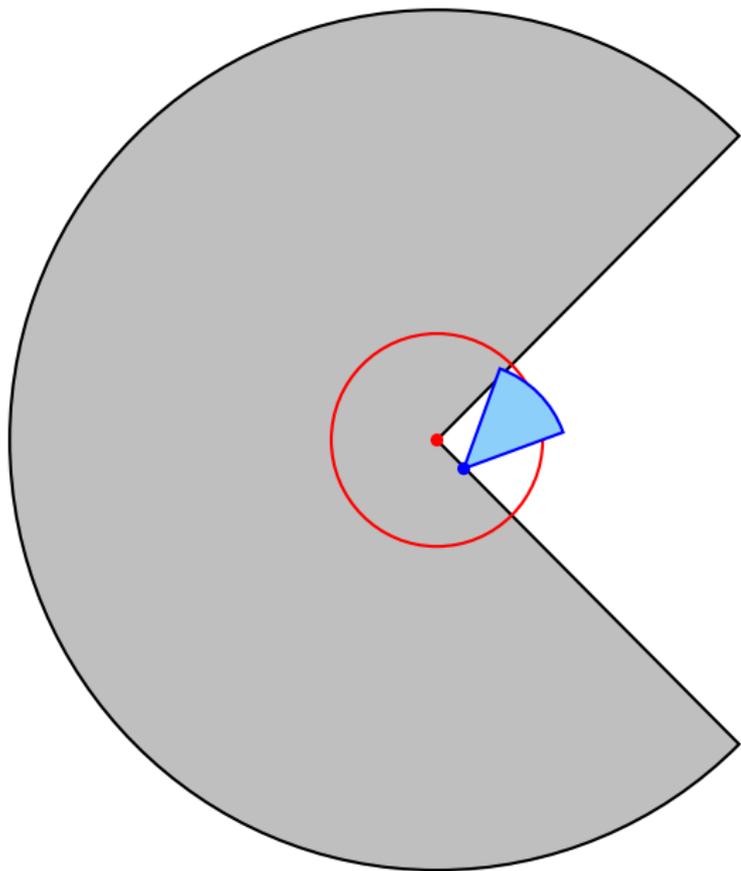
β n'est pas trop petit



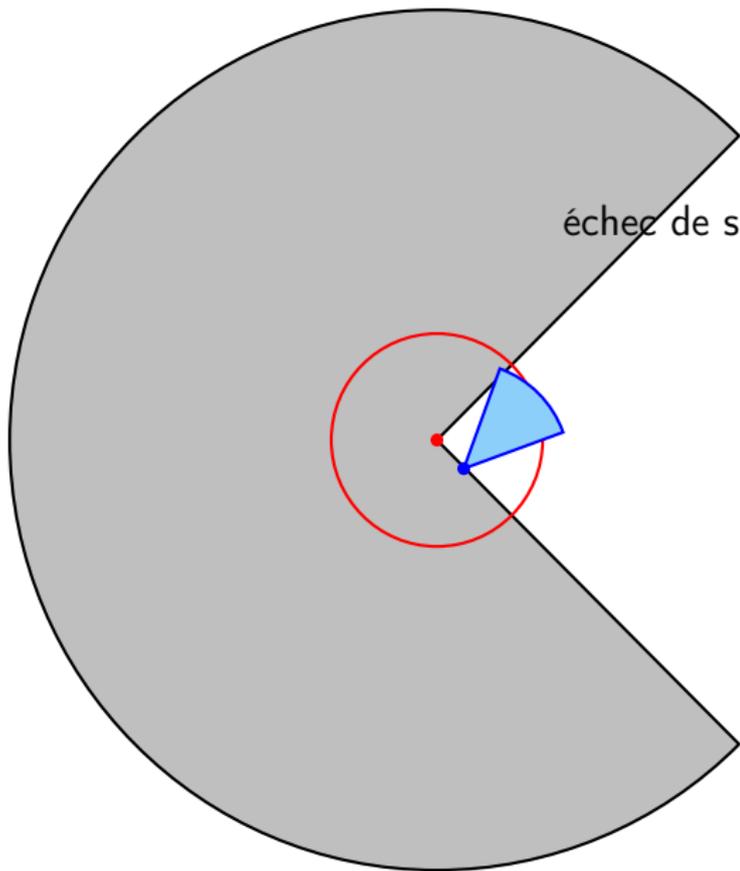
B non-convexe

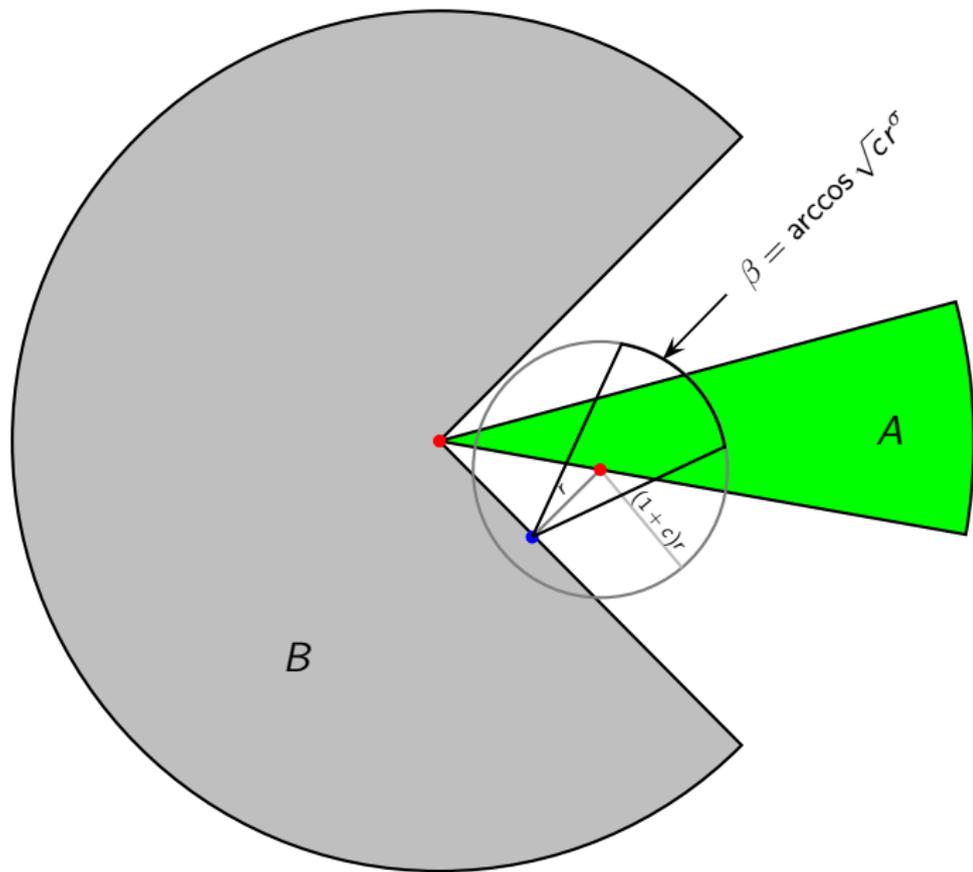
B super-régulier :

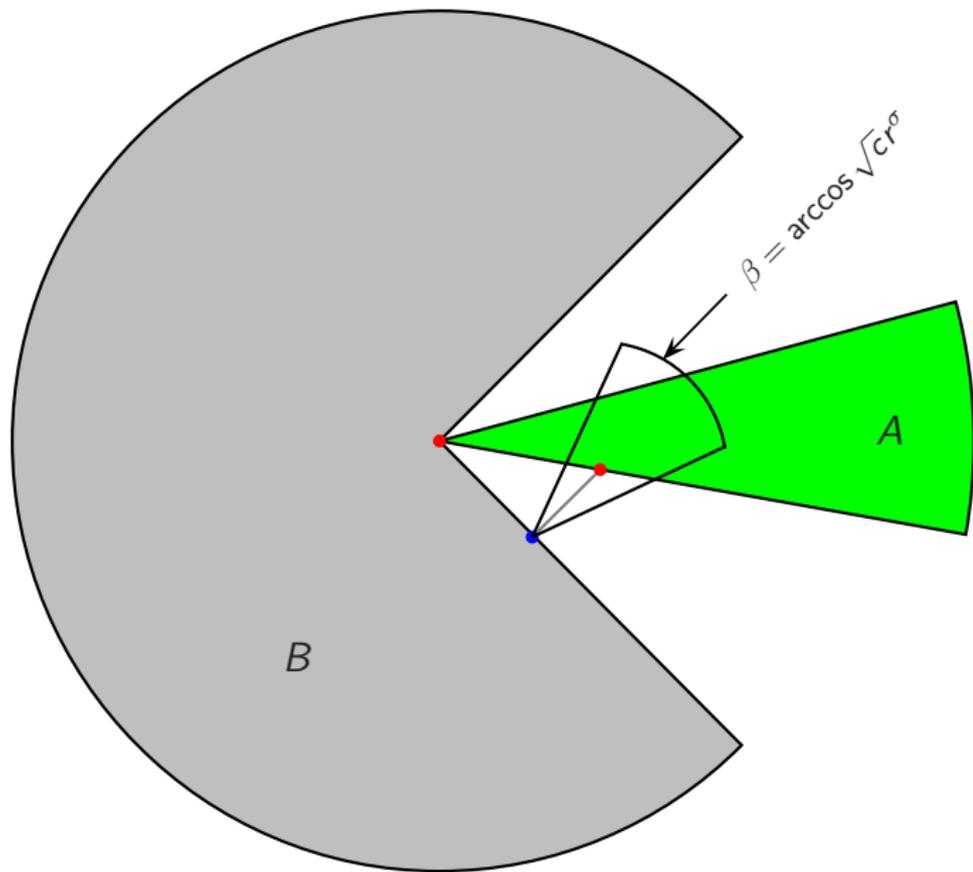
β n'est pas trop petit

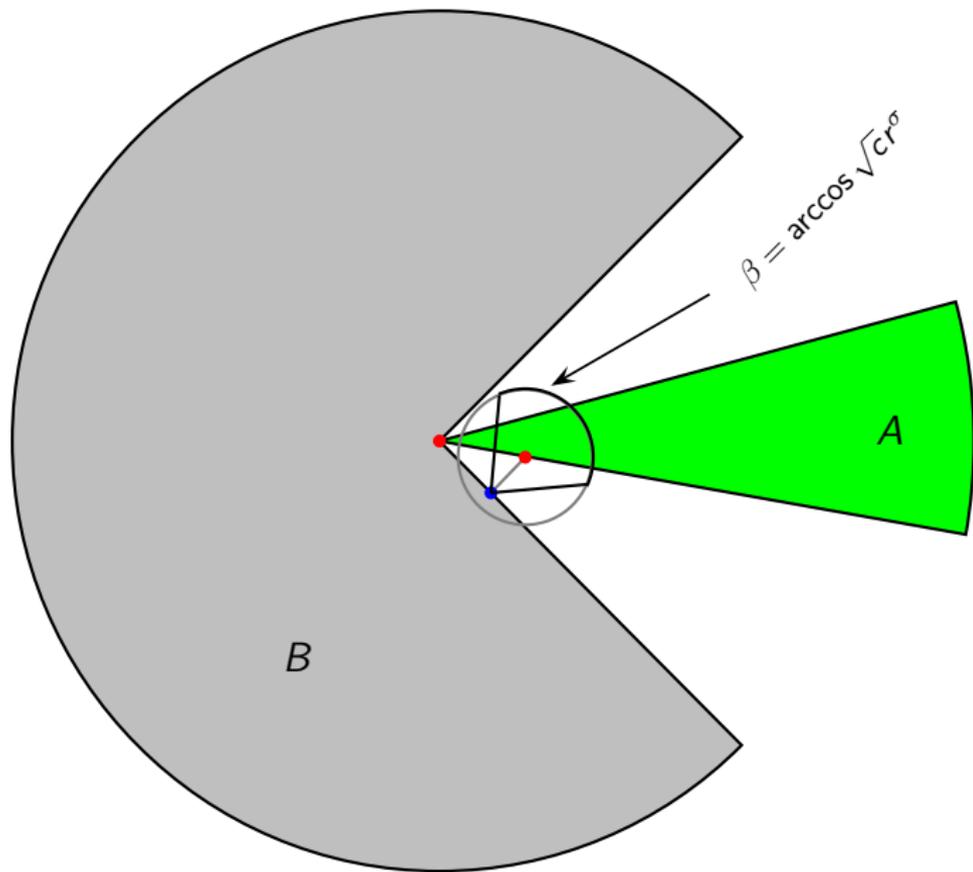


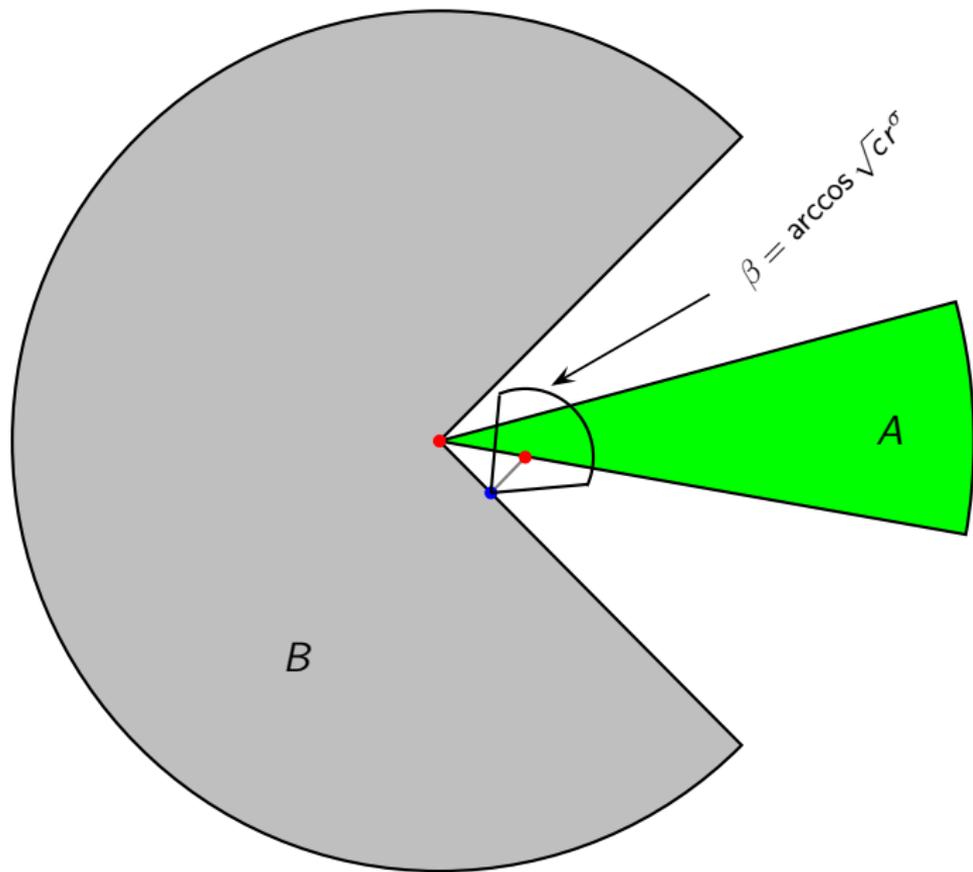
échec de super-régularité











Conséquence : notre régularité Hölderienne s'avère plus flexible.
Toujours valable pour packman.

Corollaire

(Noll, Rondepierre 2013). Soient A, B sous-analytiques, B Hölder régulier par rapport à A . Soit a_k, b_k une suite alternée bornée avec $a_k - b_k \rightarrow 0$. Alors il existe $\omega \in [0, 2)$ tel que la suite converge vers un $a^* \in A \cap B$ avec vitesse

$$\|a_k - a^*\| = \mathcal{O}\left(k^{-\frac{2-\omega}{2\omega}}\right), \quad \|b_k - a^*\| = \mathcal{O}\left(k^{-\frac{2-\omega}{2\omega}}\right)$$

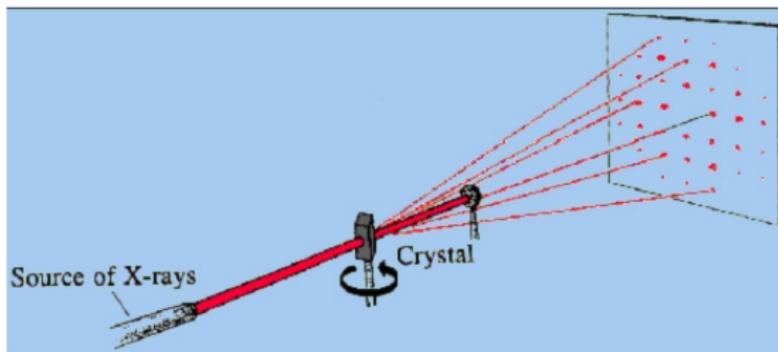
Problème de phase (Phase retrieval)

Problème de phase (en anglais: phase retrieval)

Reconstruire un signal $x(t)$, $t = 0, \dots, N - 1$ inconnu, à partir de son amplitude de Fourier $a(f) = |\hat{x}(f)|$, $f = 0, \dots, N - 1$ connue.

- Faut donc reconstruire la phase de Fourier $\hat{x}(f)/|\hat{x}(f)|$.
- **Sous-déterminé : Besoin d'ajouter de l'information.**
 - ⇒ Mesures additionnelles dans un deuxième plan de Fourier (Gerchberg-Saxton 1972).
 - ⇒ $\text{supp}(x)$ au domaine temporel connu :
 $x(t) = 0$ pour $t \notin S$ (atomicity constraint).

- Max von Laue (1912) propose d'utiliser les rayons X pour visualiser structure cristalline grâce à la diffraction.
- David Sayre (1952) montre que x peut être obtenu de $|\hat{x}|$ théoriquement si $a = |\hat{x}|$ est échantillonné 2 fois le taux de Nyquist dans chaque dimension.
 - ⇒ Mais déplore absence de méthode pour le faire.
 - ⇒ Participe au développement du 1er compilateur `fortran`.
- R.W. Gerchberg - O.W. Saxton (1972). Algorithme pour reconstruire x de $|\hat{x}|$.
- J. Miao, P. Charalambous, J. Kirz and D. Sayre, Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens, *Nature* 400, 342-344 (1999).
- 2014. Protéines individuelles et nano-cristaux sont visualisés.



Max von Laue
(photo 1929)



David Sayre
(photo 1972)



W. O. Saxton
(photo 2012)

Gerchberg-Saxton error reduction (1972)

- 1 Etant donné l'estimation actuelle x , calculer \hat{x} et «corriger» l'amplitude de Fourier $\hat{y}(f) := a(f) \frac{\hat{x}(f)}{|\hat{x}(f)|}$.
- 2 Calculer transformé de Fourier inverse y de \hat{y} , et «corriger» le domaine $x^+(t) = \begin{cases} y(t) & \text{pour } t \in S \\ 0 & \text{pour } t \notin S \end{cases}$.
- 3 Remplacer x par x^+ et boucler.

⇒ Optique, astronomie, cristallographie, nanomatériaux, ...

⇒ Aucune preuve de convergence depuis 1972.

Nous en fournissons la première.

Théorème

(Noll, Rondepierre 2013). *Supposons que le problème de reconstruction de phase a une solution $x^* \in A \cap B$. Si la contrainte du domaine temporel est représentée par un ensemble B sous-analytique, alors l'algorithme de Gerchberg-Saxton converge au voisinage de x^* vers une solution du problème de phase avec vitesse de convergence $\mathcal{O}\left(k^{-\frac{2-\omega}{2\omega}}\right)$ pour un certain $\omega \in [0, 2)$.*

Preuve. Une instance des projections alternées non-convexes :

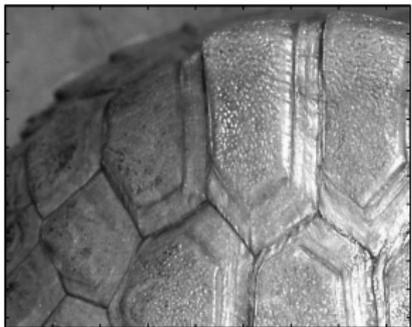
$$A = \{x \in \mathbb{C}^N : |\hat{x}(f)| = a(f) \text{ pour tout } f\}$$

A est prox-régulier et non-convexe

$$B = \{y \in \mathbb{C}^N : y(t) = 0 \text{ pour tout } t \notin S\}$$

$$P_A(x) = (a\hat{x}/|\hat{x}|)^\sim \quad P_B(y) = y \cdot \mathbf{1}_S$$





$$\hat{x} = r_t e^{i\phi_t}$$

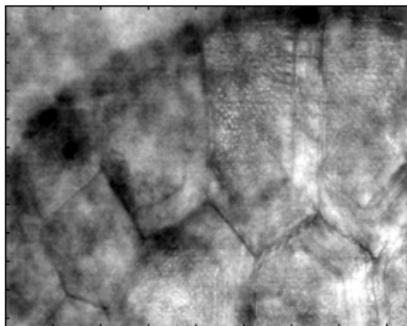


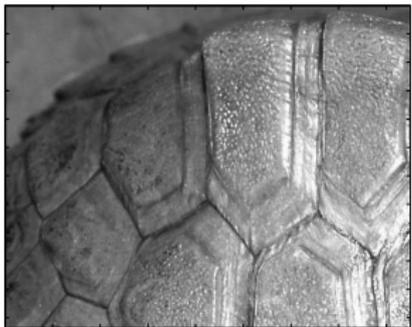
$$\hat{y} = r_g e^{i\phi_g}$$

$$r_t e^{i\phi_g}$$



$$r_g e^{i\phi_t}$$





$$\hat{x} = r_t e^{i\phi_t}$$

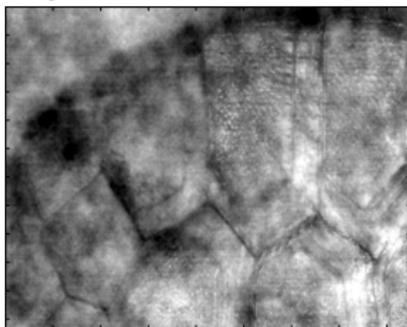


$$\hat{y} = r_g e^{i\phi_g}$$

$$r_t e^{i\phi_g}$$



$$r_g e^{i\phi_t}$$

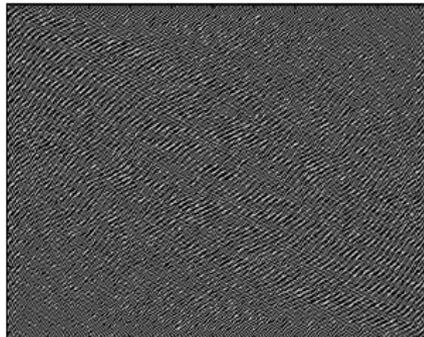


Conséquences :

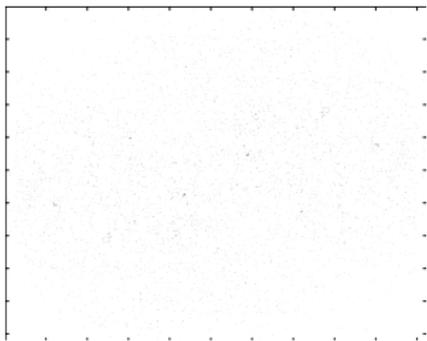
- Phase de Fourier $\hat{x}/|\hat{x}|$ donne l'information essentielle sur x .
- Amplitude de Fourier $a = |\hat{x}|$ ne permet pas de localiser l'image x .
- Exemple : traduire dans le domaine temporel change la phase de Fourier, mais non l'amplitude.
- **Reconstruction de la phase donc difficile !**



original élargi (inconnu)



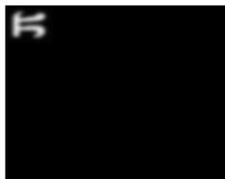
phase de Fourier (inconnue)



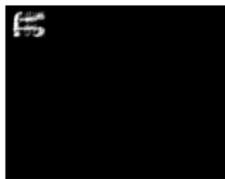
amplitude de Fourier (connue)



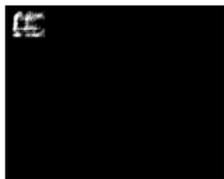
support estimé (connu a priori)



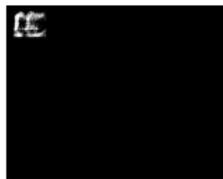
guess



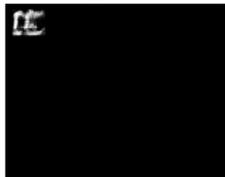
map1



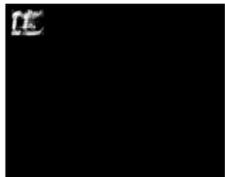
map6



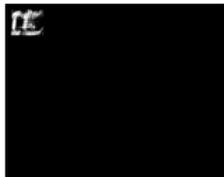
map15



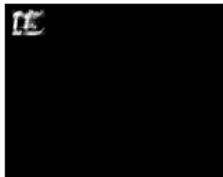
map38



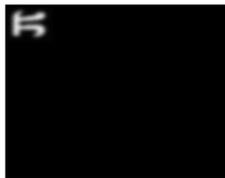
map80



map100



map200



guess



dr1



dr2



dr3



dr7



dr20



dr36



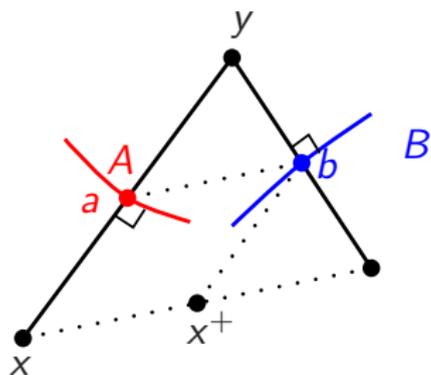
dr60

- Image idéale x_0 = PI élargie à 1024×1024 par rajout de zéros (an anglais : 0-padding).
- 0=noir, 256=blanc.
- Estimation initiale est version flouée, bruitée et tournée de 90° de Pi.
- Amplitude de Fourier $a = |\hat{x}_0|$ connue.
- $A = \{x \in \mathbb{C}^{1024 \times 1024} : |\hat{x}(f)| = a(f) \text{ pour toute fréquence } f\}$.
- $B = \{y \in \mathbb{C}^{1024 \times 1024} : y(t) = 0 \forall \text{ pixel } t \text{ hors masque}\}$.
- Masque = la zone grise autour du PI. Hypothèse a priori : masque connu.
- PA ne réussit pas complètement dans le temps donné.
- Douglas-Rachford reconstruit phase assez complètement.

J. Douglas, H.H. Rachford. On the numerical solution of heat conduction problems in two and three dimensions. *TAMS* 82 (1956), 421 – 439.

P.-L. Lions, B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM J.Num. Anal.* 16 (1979), 946–979.

J.R. Fienup. Phase retrieval algorithms : a comparison. *Applied Optics*, 1982. **HIO = Hybrid-Input-Output**



$$a \in P_A(x)$$

$$y = 2a - x \in R_A(x)$$

$$b \in P_B(y)$$

$$x^+ = x + b - a$$

reflect-reflect-average

Convergence pour les projections alternées :

A.S. Lewis, J. Malick (2008). Alternating projections on manifolds. *Math. Oper. Res.* 33, 216–234.

A.S. Lewis, R. Luke, J. Malick (2009). Local linear convergence for alternating and averaged non-convex projections. *Foundations Comp. Math.* 9, 485–513.

H.H. Bauschke, D.R. Luke, H.M. Phan, X. Wang (2013). Restricted normal cones and the method of alternating projections. *Set-valued Var. Anal.* 31, 431–501.

D. Noll, A. Rondepierre (2013). On local convergence of the method of alternating projections. *Foundations of Computational Mathematics*, 2015.

Lien avec l'inégalité de Kurdyka-Łojasiewicz :

H. Attouch, J. Bolte, P. Redont, A. Soubeyran (2010). Proximal alternating minimization and projection methods for nonconvex problems. An approach based on the Kurdyka-Łojasiewicz inequality. JMOR 35(2) :2010,438-457.

H. Attouch, J. Bolte, B.F. Svaiter (2013). Convergence of descent methods for semi-algebraic and tame problems : proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Math. Programming 137(1-2) :2013,91-129.

Convergence pour Douglas-Rachford

R. Hesse, D.R. Luke (2013). Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems, *SIAM Journal on Optimization* 23(4), 2397–2419.

H.H. Bauschke, D. Noll (2014). On the local convergence of the Douglas–Rachford algorithm. *Archiv der Mathematik* 102, 589 – 600.

H.M. Phan (2014). Linear convergence of the Douglas-Rachford method for two closed sets. *Optimization* 2015.

Comparaison de la résolution de la CDI avec :

Prix Nobel de Chimie 2014 : Microscopie fluorescente



Stefan Hell
*1962



William Moerner
*1953



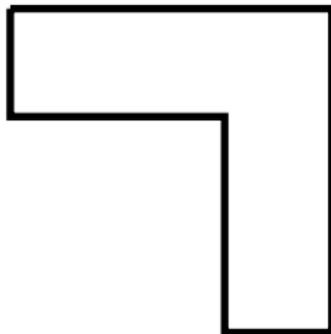
Eric Betzig
*1960

Microscopie fluorescente : $1\mu\text{m} = 10^{-6}\text{m}$

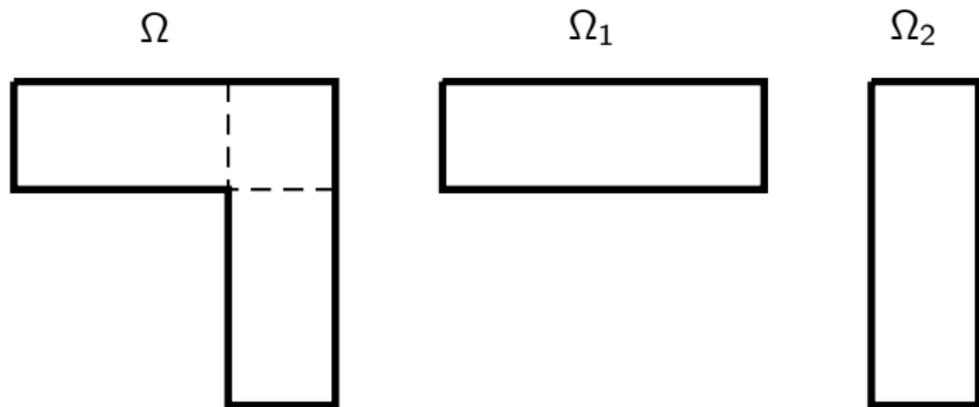
CDI : $10\text{nm} = 10^{-8}\text{m}$ (organique)

$2\text{ nm} = 2 \cdot 10^{-9}\text{m}$ (anorganique)

Merci pour votre attention !

Ω 

$$-\Delta u = f \text{ sur } \Omega \quad u = 0 \text{ sur } \partial\Omega$$



$$-\Delta u^{2k+1} = f \text{ sur } \Omega_1 \quad u^{2k+1} = u^{2k} \text{ sur } \partial\Omega_1$$

$$-\Delta u^{2k} = f \text{ sur } \Omega_2 \quad u^{2k} = u^{2k-1} \text{ sur } \partial\Omega_2$$

$$u^0 \in H_0^1(\Omega)$$

$$\langle u, v \rangle = \int_{\Omega} \nabla u \cdot \nabla v \, dx, \quad u, v \in H_0^1(\Omega)$$

$$\langle u^{2k+1} - u, v_1 \rangle = 0 \quad \forall v_1 \in H_0^1(\Omega_1), \quad u^{2k+1} - u^{2k} \in H_0^1(\Omega_1)$$

$$\langle u^{2k} - u, v_2 \rangle = 0, \quad \forall v_2 \in H_0^1(\Omega_2), \quad u^{2k} - u^{2k-1} \in H_0^1(\Omega_2)$$

$$\langle u, v \rangle = \int_{\Omega} \nabla u \cdot \nabla v \, dx, \quad u, v \in H_0^1(\Omega)$$

$$\langle u^{2k+1} - u^{2k}, v_1 \rangle = \langle u - u^{2k}, v_1 \rangle \quad \forall v_1 \in H_0^1(\Omega_1), \quad u^{2k+1} - u^{2k} \in H_0^1(\Omega_1)$$

$$\langle u^{2k} - u^{2k-1}, v_2 \rangle = \langle u - u^{2k-1}, v_2 \rangle \quad \forall v_2 \in H_0^1(\Omega_2), \quad u^{2k} - u^{2k-1} \in H_0^1(\Omega_2)$$

$$\langle u, v \rangle = \int_{\Omega} \nabla u \cdot \nabla v \, dx, \quad u, v \in H_0^1(\Omega)$$

$$u^{2k+1} - u^{2k} = P_{H_0^1(\Omega_1)}(u - u^{2k})$$

$$u^{2k} - u^{2k-1} = P_{H_0^1(\Omega_2)}(u - u^{2k-1})$$

$$u - u^{2k+1} = P_{H_0^1(\Omega_1)^\perp}(u - u^{2k})$$

$$u - u^{2k} = P_{H_0^1(\Omega_2)^\perp}(u - u^{2k-1})$$

Corollaire

Si $\Omega = \Omega_1 \cup \Omega_2$, alors la suite des projections alternées $u - u^{2k+1}$, $u - u^{2k}$ converge vers 0 en norme $H_0^1(\Omega)$.

Preuve. Il suffit de montrer que $H_0^1(\Omega_1) + H_0^1(\Omega_2)$ est dense dans $H_0^1(\Omega)$. Pour cela il suffit de montrer que tout $\phi \in \mathcal{D}(\Omega)$ peut être écrit comme $\phi = \phi_1 + \phi_2$ avec $\phi_i \in \mathcal{D}(\Omega_i)$. □

Application : Algorithme EM

A.P. Dempster, N.M. Laird, D.B. Rubin. Maximum likelihood from incomplete data via the EM-algorithm.

J. Royal Stat. Soc. Series B, vol. 39, no. 1 (1977), 1 – 38.

⇒ Cité 38230 fois depuis 1977

⇒ Hélas, pas de preuve de convergence.

⇒ Quelques preuves dans les situations particulières convexes.

Notre méthode donne une preuve de convergence locale dans le cas d'une loi normale avec domaine paramétrique non-convexe.

Structured low-rank matrix approximation

Given a structured matrix $x \in \mathcal{S}$, solve the problem

$$(P) \quad \begin{array}{ll} \text{minimize} & \|x' - x\|_F \\ \text{subject to} & x' \in \mathcal{S} \\ & \text{rank}(x') \leq r \end{array}$$

- \mathcal{S} = Hankel matrices (denoising of time series)
- \mathcal{S} = Toeplitz matrices (spectral estimation problems)
- \mathcal{S} = positive semidefinite matrices
- \mathcal{S} = stable matrices

Use non-convex alternating projections :

$$A = \{x : x \in \mathcal{S}\} \quad P_A = \text{easy ???}$$

$$B = \{x : \text{rank}(x) \leq r\} \quad P_B = \text{truncated SVD}$$

Can now prove local convergence to $x^* \in \{x : \text{rank}(x) \leq r, x \in \mathcal{S}\}$. Need not be solution of (P)

Sparse affine feasibility

$$\begin{aligned}
 &\text{minimize} && \|x\|_0 = \text{number of non-zero entries in } x \\
 &\text{subject to} && Ax = b \\
 &&& x \in \mathbb{R}^n
 \end{aligned}$$

Use non-convex alternating projections :

$$A = \{x \in \mathbb{R}^n : \|x\|_0 \leq k\} = \bigcup_{\text{card}(I) \leq k} \underbrace{\text{span}\{e_i : i \in I\}}_{=: A_I}$$

$$P_A(x) = \bigcup_{I_{\text{active}}} P_{A_I}(x)$$

$$B = \{x \in \mathbb{R}^n : Ax = b\} \quad P_B(x) = x - A^\dagger(Ax - b)$$

Packings in Grassmannian manifolds

$\mathbb{G}(k, \mathbb{C}^d) =$ all k -dimensional subspaces of \mathbb{C}^d

Represent $\underline{S} \in \mathbb{G}(k, \mathbb{C}^d)$ by unitary $S \in \mathbb{C}^{k \times d} : S^* S = I_k, \text{range}(S) = \underline{S}$

For two subspaces S, T do SVD :

$$S^* T = UCV^*$$

then $c_{kk} = \cos \theta_k$ the principal angles between $\underline{S}, \underline{T}$. Leads to distances between \underline{S} and \underline{T} :

- Chordal distance : $\sqrt{\sin^2 \theta_1 + \dots + \sin^2 \theta_k} = (k - \|S^* T\|_F^2)^{1/2}$
- Spectral distance : $\min_i \sin \theta_i = (k - \|S^* T\|_{2,2}^2)^{1/2}$
- Fubini-Study distance : $\arccos(\prod_j \cos \theta_j)$
- Geodesic distance : $\sqrt{\theta_1^2 + \dots + \theta_k^2}$

Packing for the chordal distance :

$$\text{pack}(S_1, \dots, S_N) := \min_{m \neq n} d_{\text{chord}}(S_m, S_n) = \min_{m \neq n} (k - \|S_m^* S_n\|_F^2)^{1/2}$$

True problem : $\max_{\{S_1, \dots, S_N\}} \text{pack}_{\text{chord}}(S_1, \dots, S_N)$

Instead feasibility problem : Given $\rho > 0$, want $\{S_1, \dots, S_N\}$ such that $\text{pack}_{\text{chord}}(S_1, \dots, S_N) \geq \rho$

$$\min_{m \neq n} (k - \|S_m^* S_n\|_F^2)^{1/2} \geq \rho \equiv \max_{m \neq n} \|S_m^* S_n\|_F \leq \mu := \sqrt{k - \rho^2}$$

Put $S := [S_1 S_2 \dots S_N]$

Gramian $G = S^* S \in \mathbb{C}^{kN \times kN} \succeq 0$, $G_{mn} = S_m^* S_n$

- 1 G is Hermitian
- 2 Each diagonal block of G is identity I_k
- 3 $G \succeq 0$
- 4 $\text{rank}(G) \leq d$
- 5 $\text{trace}(G) = kN$

Conversely, any G with these properties can be factored $G = S^* S$ and $S = [S_1 \dots S_N]$ gives then rise to a configuration of N subspaces in the Grassmannian $\mathbb{G}(k, \mathbb{C}^d)$.

Now alternating projections :

Structural constraint (convex) :

$$A = \{H \in \mathbb{C}^{kN \times kN} : H = H^*, H_{nn} = I_k, \|H_{mn}\|_F \leq \mu\}$$

Spectral constraint (non-convex) :

$$B = \{G \in \mathbb{C}^{kN \times kN} : G \succeq 0, \text{rank}(G) \leq d, \text{trace}(G) = kN\}$$

Any solution $G \in A \cap B$ gives a packing of size N of the Grassmannian with chordal distance packing index $\geq \rho$.

Compute projection on A (easy) :

$$H = P_A(G) : H_{mn} = \begin{cases} G_{mn} & \|G_{mn}\|_F \leq \mu \\ \mu G_{mn} / \|G_{mn}\|_F & \text{else} \end{cases}$$

Compute projection on B (more involved but possible) :

Let $H = \sum_{j=1}^{kN} \lambda_j u_j u_j^*$ be spectral decomposition, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{kN}$. Then

$$G = \sum_{j=1}^d (\lambda_j - \gamma)_+ u_j u_j^* \in P_B(H)$$

provided γ is chosen such that $\sum_{j=1}^d (\lambda_j - \gamma)_+ = kN$.