N° d'ordre : 1371

THÈSE

présentée à

L'UNIVERSITÉ DES SCIENCES ET TECHNOLOGIES DE LILLE

Pour obtenir

LE TITRE DE DOCTEUR DE L'UNIVERSITÉ

SPÉCIALITÉ : MATHÉMATIQUES PURES

Par

Abderrahim MAKKI NACIRI

Approximation par son cône tangent d'une partie définie par des égalites et des inégalités

Soutenue le 27 Septembre 1994 devant la Commission d'Examen :

Président : G.COEURÉ, Université de Lille
Rapporteurs : J.BAIR , Université de Liège
S.NICAISE , Université de Valenciennes
Examinateurs : Ph.ANTOINE , Université de Lille
M.MBEKHTA , Université de Lille

Je tiens à remercier Monsieur Philippe Antoine qui m'a proposé le sujet de cette thèse et qui m'a bien guidé durant la préparation de ce travail. Avec lui, j'ai appris beaucoup de choses ; qu'il veuille bien trouver ici l'expression de ma profonde reconnaissance.

Je remercie Monsieur le Professeur Gérard Cœuré d'avoir bien voulu présider le jury de cette thèse, ainsi que Messieurs Jacques Bair et Serge Nicaise qui ont accepté de juger ce travail.

Je remercie également Monsieur le Professeur Mostafa Mbekhta pour sa gentillesse et d'avoir accepté de faire partie du jury.

A mes parents.

...A mes frères et soeurs.

SOMMAIRE

Introduction

Soient E un espace de Banach, W une partie de E et a un point de W. Si W est définie d'une manière implicite, par exemple comme ensemble des solutions d'un système d'équations et d'inéquations, et si V est un voisinage de a, il n'est pas toujours possible de déterminer explicitement une région de E qui contienne d'une manière significative la trace de W sur V, ni une région ne la rencontrant pas. Une approche possible consiste à se ramener à une approximation convenable de W au voisinage de a. Rappelons une définition classique :

<u>Définition</u> : On appelle <u>cône tangent</u> en a à W et on note T_aW, l'ensemble des éléments x de E tels qu'il existe une suite $(x_n)_{n\in\mathbb{N}}$ de W tendant vers a et une suite $(\alpha_n)_{n\in\mathbb{N}}$ de $\mathbb{R}^{+*}$ telles que $\alpha_n^{-1}(x_n - a)$ tende vers x. (C'est le cône contingent au sens de Bouligand). On verifie aisément que ce cône est fermé.

La détermination du cône tangent en a à W permet une étude infinitéimale de W au voisinage de a. C'est par exemple une voie classique pour établir une condition nécessaire d'optimalité en a sur W pour une fonctionnelle différentiable. Plus généralement la connaissance de T_aW permet une exploration qualitative de W au voisinage de a :

. Si X n'est pas un vecteur tangent en a à W, il existe un voisinage U de X et un voisinage V de a tels que le cône pointé $a + \mathbb{R}^{+*}.U$ ne rencontre pas W dans V.

. Si X est un vecteur tangent en a à W, quels que soient les voisinages U de X et V de a, le cône pointé $a + \mathbb{R}^{+*}.U$ rencontre W dans V.

Ainsi [1] - en déterminant le cône tangent à une partie définie par des égalités et des inégalités - a obtenu des renseignements sur l'ensemble des solutions d'un problème d'optimisation paramètré.

Les limites de cette exploration de W sont de deux ordres:

. Dans le cas d'un vecteur non tangent, l'aspet quantitatif est absent : on a aucun contrôle sur la taille des voisinages U et V

. Dans le cas d'un vecteur tangent, l'insuffisance tient à la formulation même du resultat : On exprime en effet que toute direction dans le cône tangent est limite de directions de points de W alors qu'on souhaiterait savoir si W est "approximé" par T_aW.

Nous nous proposons ici de faire plus généralement une étude de l'approximation de W par un cône fermé a priori différent de T_aW. Au chapitre I, nous étudierons des généralites sur cette approximation, au chapitre II nous aborderons l'aspect quantitatif de cette approximation dans le cas où W est définie par des égalités et des inégalités, et au chapitre III nous appliquons ce resultat à un problème d'optimisation paramètré pour obtenir des renseignements quantitatifs sur la variation de l'ensemble des solutions au voisinage d'un paramètre donné.

CHAPITRE.I

APPROXIMATION LOCALE D'UNE PARTIE PAR UN CÔNE FERMÉ

1.1. Notion d'approximation locale par un cône fermé.

Soient E un espace de Banach, W une partie de E et a un point de W. Nous noterons W^* l'ensemble $W - \{a\}$.

Définition 1.1.1 ([2]) : Soit $Q \subset E$ un cône fermé. On dit que W est approximé par Q au point a si :

(1): $\forall \varepsilon > 0, \exists \eta > 0, \forall x \in W^* \cap \bar{B}(a,\eta), \exists X \in Q, \|x - a - X\| < \varepsilon . \|x - a\|$.

Cette définition exprime que $W^* \cap \bar{B}(a,\eta)$ est contenu dans $a + Q_\varepsilon$ où Q_ε est l'"épaississement" conique de Q défini par :

$$Q_\varepsilon = \{ X \in E \,/\, d(X, Q) < \varepsilon \|X\| \}.$$

Elle permet donc de localiser explicitement W au voisinage de a pourvu qu'on sache déterminer explicitement η en fonction de ε.

Remarque 1.1.2 : Les notations sont les mêmes. Si on connaît explicitement une application surjective, strictement nonotone $\varepsilon \longmapsto r(\varepsilon)$ de $\mathbb{R}^{+*}$ dans $\mathbb{R}^{+*}$ telle que pour tout $\varepsilon > 0$, la propriété (1) est verifieé en prenant $\eta = r(\varepsilon)$, alors en notant $\eta \longmapsto e(\eta)$ l'application réciproque de r ona la localisation suivante:

$$(2): \forall \eta > 0, \quad W^* \cap \bar{B}(a, \eta) \subset a + Q_{e(\eta)},$$

qui exprime que pour tout $\eta > 0$, la trace de W sur $\bar{B}(a, \eta)$ est contenue dans la région explicite $a + Q_{e(\eta)}$.

Cette formulation peut être interpretée de manière globale : Si on note $\tilde{Q}$ le "horn_neighborhood" de Q défini par :

$$\tilde{Q} = \{ x \in E, \exists x' \in Q, \|x - x'\| < e(\|x\|) . \|x\| \},$$

alors (2) entraîne d'une manière simple l'inclusion globale :

$$W^* \subset a + \tilde{Q} \quad ,$$

qui est localement plus significative que (2), car la restriction de $a + \tilde{Q}$ à toute boule de centre a et de rayon η est strictement incluse dans $a + Q_{e(\eta)}$.

<u>Exemple</u> : Dans $\mathbb{R}^2$, soient W le cercle de centre o

et de rayon R, a le point $(0,R)$ et $Q = T_a W$, c'est à dire l'axe des x. Soit $\varepsilon > 0$. Sur le schéma (voir page), $a + Q_\varepsilon$ est le cône de sommet a, situé strictement entre les deux droites Q_1 et Q_2. Par un calcul trigonométrique simple on trouve que : $r = 2R.\varepsilon$. Donc pour tout $c < 2R$ on a l'inclusion suivante :

$$\bar{B}(a, c.\varepsilon) \cap W^* \subset a + Q_\varepsilon,$$

ce qui traduit que W est approximé au point a par le cône Q. Pour mesurer cette approximation, il suffira de prendre : $r(\varepsilon) = c.\varepsilon$ avec $c < 2R$. Il vient alors si on pose $e(\eta) = \frac{\eta}{c}$:

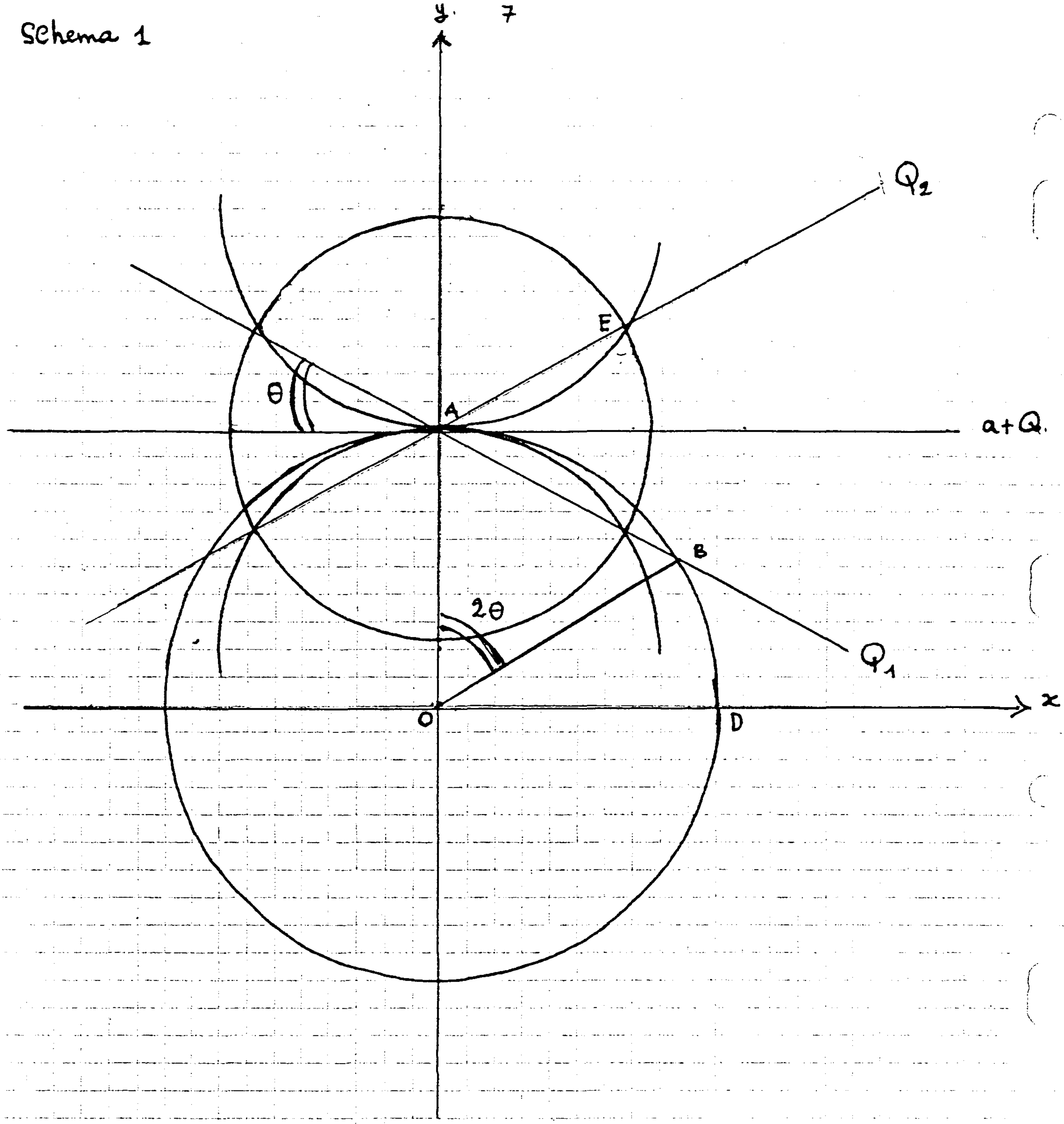
$$\forall \eta > 0, \quad W^* \cap \bar{B}(a, \eta) \subset a + Q_{\frac{\eta}{c}}.$$

Le horn-neighborhood de Q est explicite et il est égal à :

$$\tilde{Q} = \left\{ (x,y) \in \mathbb{R}^2, \ |y| < \frac{x^2+y^2}{c} \right\}.$$

$a + \tilde{Q}$ est donc le complementaire des deux disques de centres $(0, R - \frac{c}{2})$ et, $(0, R + \frac{c}{2})$ et de rayons égaux à $\frac{c}{2}$.

Schema 1



- $\sin\theta = \varepsilon$
- $d(A,B) = r = 2R.\varepsilon$
- $d(A,E) = C.\varepsilon$
- $d(O,D) = R$

1.2 Recherche du meilleur approximant

Dans la pratique, il est évident qe la qualité de l'approximation dépend du cône avec lequel on approxime et surtout de sa taille : Si Q_1 et Q_2 sont deux cônes fermés qui approximent W au point a et tels que Q_1 est inclus dans Q_2 alors l'approximation de W par Q_1 est plus significative que celle de W par Q_2.

On a le resultat suivant :

Lemme 1.2.1 : Si Q est un cône fermé qui approxime W au point a alors il contient $T_a W$.

Démonstration : Soit X un élément de $T_a W$:
Il existe une suite $(x_n)_{n\in\mathbb{N}}$ de W tendant vers a et une suite $(\alpha_n)_{n\in\mathbb{N}}$ de $\mathbb{R}^{+*}$ telles que $\alpha_n^{-1}(x_n - a)$ tende vers X.

Soit $\varepsilon > 0$. W est approximé par Q au point a, donc :

$$(3) : \exists \eta > 0, \forall x \in W \overset{*}{\cap} \bar{B}(a,\eta), \exists X' \in Q, \|x - a - X'\| < \varepsilon . \|x - a\|.$$

La limite de $(x_n)_{n\in\mathbb{N}}$ étant égale à a, il existe un entier N tel que pour tout $n \geq N$ on ait $\|x_n - a\| < \eta$.
Il vient donc de (3) :

$$\forall n \geqslant N,\ \exists X_n' \in Q,\ \| x_n - a - X_n' \| < \varepsilon . \| x_n - a \|,$$

et si on pose : $X_n = \alpha_n^{-1}(x_n - a)$ et $X_n'' = \alpha_n^{-1} . X_n'$:

$$(4) : \forall n \geqslant N,\ \| X_n - X_n'' \| < \varepsilon . \| X_n \| .$$

La suite $(X_n)_{n \in \mathbb{N}}$ étant convergente (vers X), il résulte que $(X_n'')_{n \in \mathbb{N}}$ est une suite de cauchy, donc converge vers un élément X'', et comme $(X_n'')_{n \in \mathbb{N}}$ est une suite d'éléments du fermé Q, on a $X'' \in Q$. Si on passe donc à la limite dans (4), il vient :

$$\| X - X'' \| \leqslant \varepsilon . \| X \| .$$

ε est quelconque, donc $X = X''$ et par conséquent X appartient à Q. Ce qui montre que $T_a W$ est inclus dans Q ∎

Il résulte de ce lemme que $T_a W$ est un minorant au sens de l'inclusion des cônes fermés qui approximent W au point a. Donc si W est approximé par $T_a W$ au point a alors cette approximation est optimale.

<u>Lemme</u> 1.2.2 : Si $\dim E < \infty$, alors W est approximé par $T_a W$ au point a.

Démonstration : Si cela était faux, alors :

$$\exists \varepsilon>0, \forall \eta>0, \exists x \in W^* \cap \bar{B}(a,\eta), \forall x' \in T_a W, \|x-a-x'\| \geqslant \varepsilon.\|x-a\|.$$

En posant pour tout $n \in \mathbb{N}^*$, $\eta = \frac{1}{n}$, il vient donc :

(5): $\exists \varepsilon>0, \forall n \in \mathbb{N}^*, \exists x_n \in W^*, \|x_n-a\| \leqslant \frac{1}{n}, \forall x' \in T_a W, \|x_n-a-x'\| \geqslant \varepsilon.\|x_n-a\|.$

On pose $X_n = \frac{x_n-a}{\|x_n-a\|}$. La sphère unité étant compacte en dimension finie, il existe une sous-suite $X_{\varphi(n)} = \frac{x_{\varphi(n)}-a}{\|x_{\varphi(n)}-a\|}$ de $(X_n)_{n\in\mathbb{N}}$ qui converge vers un élément X, et comme $x_{\varphi(n)}$ tend vers a, X appartient au cône T_aW et pour tout $n \in \mathbb{N}^*$, $\|x_{\varphi(n)}-a\|.X$ appartient aussi à T_aW. En prenant donc $X' = \|x_{\varphi(n)}-a\|.X$ dans (5), il vient :

$$\forall n \in \mathbb{N}^*, \ \varepsilon.\|x_{\varphi(n)}-a\| \leqslant \|(x_{\varphi(n)}-a) - \|x_{\varphi(n)}-a\|.X\|.$$

Ce qui entraîne si on divise par $\|x_{\varphi(n)}-a\|$ et si on passe à la limite :

$$\varepsilon \leqslant \|X-X\| = 0.$$

Ceci est impossible car ε est strictement positif ■

Cependant en dimension infinie, W n'est pas toujours approximé par $T_a W$.

<u>Exemple</u> : Sur l'ensemble ℓ^2 des suites de carrés sommables, on définit la fonctionnelle f par :

$$f(x_*) = \sum_{i=1}^{\infty} \frac{1}{n} x_n^2 - \sum_{n=1}^{\infty} x_n^3 \qquad , \text{où } x_* = (x_n)_{n\in\mathbb{N}^*} .$$

Soit $W = f^{-1}(f(0_*))$ la surface de niveau passant par 0_*. Alors W n'est pas approximé par $T_{0_*}W$ au point 0_*. D'après [3], $T_{0_*}W = \{0\}$, donc si W est approximé par $T_{0_*}W$ alors 0_* est un point isolé de W. En effet, en prenant $\varepsilon = \frac{1}{2}$ dans la définition 1.1.1, on a :

$$\exists\, \eta > 0 , (x_* \in W^*, \|x_*\| \leq \eta) \Rightarrow (\|x_*\| < \tfrac{1}{2}\|x_*\|) ,$$

Ce qui entraîne que : $W \cap B(0_*, \eta) = \{0_*\}$.

Il suffit donc de montrer que 0_* n'est pas un point isolé de W pour conclure. Soit $(e_*^p)_{p\in\mathbb{N}^*}$ la suite d'éléments de W définie par :

$$e_n^p = \begin{cases} 0 & \text{si } n \neq p \\ \frac{1}{p} & \text{si } n = p \end{cases}$$

La suite $(e_*^p)_{p\in\mathbb{N}^*}$ converge dans ℓ^2 vers O_* car pour tout p de $\mathbb{N}^*$, $\|e_*^p\| = \frac{1}{p}$. Donc O_* n'est pas un point isolé de W et par conséquent W n'est pas approximé par $T_{O_*}W$ au point O_* ∎

Le fait que W soit approximé par son cône tangent en a n'est pas suffisant pour localiser W au voisinage de a, car même si cette approximation peut être prouvée théoriquement (comme dans le cas de dimension finie) elle est souvent non efficace pour deux raisons:

- On ne sait pas en général déterminer explicitement ce cône tangent.
- Même si on connaît le cône tangent, on ne sait pas mesurer l'approximation (i.e expliciter η en fonction de ε dans la définition 1.1.1).

Ces difficultés nous ont amené à chercher à approximer W par un cône fermé explicitement déterminé, différant aussi peu que possible du cône tangent et pour lequel on sache mesurer l'approximation. C'est ce projet qui a motivé l'introduction de la notion d'approximation par un cône fermé quelconque.

CHAPITRE. II
APPROXIMATION D'UNE PARTIE DÉFINIE PAR DES ÉGALITÉS ET DES INÉGALITÉS. ([4])

2.1. Notations et hypothèses

Si n est un entier positif, nous noterons $[n]$ l'ensemble $\{1,2,\dots n\}$, et si I est une partie de $[n]$, nous noterons $|I|$ le cardinal de I, $\mathbb{R}^n_I$ le sous-espace vectoriel de $\mathbb{R}^n$ défini par :

$$\mathbb{R}^n_I = \{ X \in \mathbb{R}^n \ / \ \forall i \in I,\ X_i = 0 \},$$

et $\mathbb{R}^{n+}_I$ le cône polyédral convexe fermé de $\mathbb{R}^n$ défini par:

$$\mathbb{R}^{n+}_I = \{ X \in \mathbb{R}^n / \ \forall i \in I,\ X_i = 0 \text{ et } \forall j \in [n],\ X_j \geqslant 0 \}.$$

Si u est une application linéaire continue d'un Banach E dans $\mathbb{R}^n$, nous noterons $\mu(u)$ la norme de l'inverse de l'isomorphisme induit par u de $E/_{\operatorname{Ker} u}$ dans $\operatorname{Im} u$, et si I est une partie de $[n]$ qui s'écrit $I = \{i_1, i_2, \dots i_k\}$ avec $i_1 < i_2, \dots < i_k$, nous noterons u_I l'application linéaire de E dans $\mathbb{R}^k$ définie par :

$$u_I : E \longrightarrow \mathbb{R}^k,\quad x \longmapsto (u_{i_1}(x), u_{i_2}(x), \dots, u_{i_k}(x)).$$

Soient Ω un ouvert de l'espace de Banach E, (f,g) une application de classe C^1 de Ω dans $\mathbb{R}^p \times \mathbb{R}^q$, W la partie de Ω définie par :

$$W = \{ x \in \Omega \; / \; f(x) = 0 \text{ et } \forall i \in [q] \, , \; g_i(x) \geqslant 0 \} ,$$

et a un point de W. On pose $I_0 = \{ i \in [q] \, , \; g_i(a) = 0 \}$, $g_0 = (g_i)_{i \in I_0}$, $m = p + |I_0|$ et $h = (f, g_0)$. L'application h prend ses valeurs dans $\mathbb{R}^p \times \mathbb{R}^{|I_0|}$ que nous identifierons à $\mathbb{R}^m$..

<u>Définition 2.1.1</u> : On appelera cône des conditions d'ordre 1 de W au point a et on notera Q_a, l'ensemble

$$Q_a = Dh(a)^{-1} \left(\mathbb{R}^{m+}_{[p]} \right) = \{ X \in E \; / \; Df(a).X = 0 \text{ et } \forall i \in I_0 , \; Dg_i(a).X \geqslant 0 \} .$$

Le cône tangent en a à W est toujours inclus dans le cône Q_a et si la condition de normalité :

$$\operatorname{Im} Dh(a) - \mathbb{R}^{m+}_{[p]} = \mathbb{R}^m$$

est vérifiée, alors $T_a W = Q_a$. ([5]) .

Le théorème suivant établit l'approximation de W par Q_a au point a , et donne une mesure de cette approximation sous l'hypothèse que Dh vérifie une condition de croissance au point a. Si la condition de normalité est vérifiée, l'approximation est optimale, sinon elle reste significative tant que Q_a n'est pas trop différent de T_aW.

<u>Théorème d'approximation</u> : Les hypothèses et les notations sont celles qui ont été précisées dans les paragraphes précédents. On suppose que Ω est étoilé au point a et qu'il existe $A>0$ telle que :

$$\forall x \in \Omega - \{a\}, \quad \|Dh(x) - Dh(a)\| < A . \|x - a\| .$$

Alors :

1) Pour toute partie I de $[m]$, il existe une constante $N_I \geqslant 1$ telle que :

$$\forall x \in \mathbb{R}^m_I, \quad d(x, \mathbb{R}^m_I \cap \operatorname{Im} Dh(a)) \leqslant N_I . d(x, \operatorname{Im} Dh(a))$$

2) W est approximé par Q_a au point a et on a :

$$\forall \varepsilon > 0, \quad W^* \cap \overline{B}\left(a, \frac{2\varepsilon}{M.A}\right) \subset a + (Q_a)_\varepsilon ,$$

où $$M = \mu(Dh(a)) . \left[\sup_{[p] \subset I \subset [m]} N_I\right]^{|I_0|} .$$

Ce théorème exprime que W est approximé au point a par le cône fermé Q_a. De plus une application $\varepsilon \longmapsto r(\varepsilon)$ qui permet de mesurer l'approximation est explicite pourvu qu'on sache calculer la constante M : il suffit de prendre $r(\varepsilon) = c \cdot \varepsilon$ avec $c \leqslant \frac{2}{M \cdot A}$.

Nous allons dans une première partie établir un groupe de resultats géometriques dans $\mathbb{R}^m$, puis dans une deuxième partie démontrer le théorème d'approximation et enfin dans une troisième partie évaluer la constante M.

2.2 Quelques propriétés métriques de $\mathbb{R}^m$.

Notations : Si x est un point de $\mathbb{R}^m$, nous noterons $\langle x \rangle$ la droite engendreé par x, $\langle x \rangle^+ = \{\alpha . x \,/\, \alpha \in \mathbb{R}^+\}$ la demi-droite positive engendreé par x et si B est une partie non vide de $\mathbb{R}^m$, $d(x, B) = \inf\limits_{y \in B} \| x - y \|$ la distance euclidienne de x à B.

Rappel 2.2.1 : Si B_1 et B_2 sont deux parties non vides de $\mathbb{R}^m$ telles que B_1 est incluse dans B_2 et si x est un point de $\mathbb{R}^m$, alors $d(x, B_2) \leqslant d(x, B_1)$, et s'il existe un point y de B_1 tel que $d(x, B_2) = \| x - y \|$ alors $d(x, B_2) = d(x, B_1)$.

Lemme de réciprocité : Dans $\mathbb{R}^m$, soient K un cône convexe fermé et H un cône fermé. S'il existe une constante $N \geq 1$ telle que :

$$\forall y \in H,\ d(y, H \cap K) \leq N.d(y, K)$$

alors on a aussi :

$$\forall x \in K,\ d(x, H \cap K) \leq N.d(x, H).$$

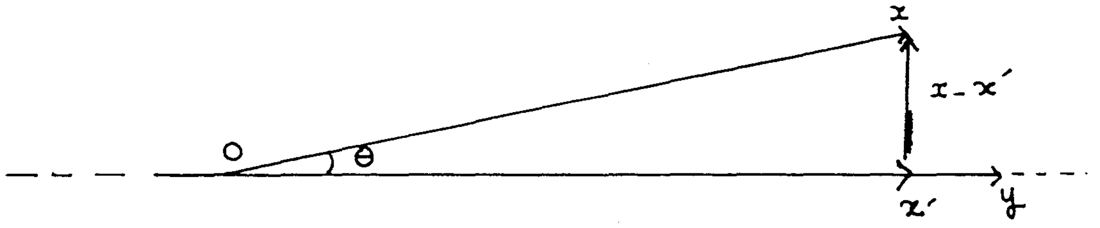
Démonstration : Nous allons d'abord établir une série de sous_lemme utiles pour la démonstration.

Lemme 2.2.2 : Si x et y sont deux points de $\mathbb{R}^m$, alors:

$$\text{a) } d(x, \langle y \rangle) = \begin{cases} \|x\|, & \text{si } y = 0 \\ \left(\|x\|^2 - \dfrac{\langle x, y\rangle^2}{\|y\|^2} \right)^{1/2}, & \text{si } y \neq 0. \end{cases}$$

$$\text{b) } d(x, \langle y \rangle^+) = \begin{cases} \|x\|, & \text{si } \langle x, y \rangle \leq 0 \\ d(x, \langle y \rangle), & \text{si } \langle x, y \rangle > 0. \end{cases}$$

Démonstration : a) Si $y = 0$, alors $d(x, \langle y \rangle) = d(x, \{0\}) = \|x\|$.
Si $y \neq 0$. Soit x' la projection de x sur $\langle y \rangle$.



On a : $\cos\theta = \dfrac{\langle x,y\rangle}{\|x\|\cdot\|y\|} = \dfrac{\langle x,x'\rangle}{\|x\|\cdot\|x'\|}$ et $\langle x-x', x'\rangle = 0$.

Donc : $\|x'\| = \dfrac{\langle x,y\rangle}{\|y\|}$. Si on applique le théorème de Pythagore au triangle (o, x, x') il vient alors :

$$d^2(x,\langle y\rangle) = \|x-x'\|^2 = \|x\|^2 - \|x'\|^2 = \|x\|^2 - \frac{\langle x,y\rangle^2}{\|y\|^2}$$

soit :

$$d(x,\langle y\rangle) = \left(\|x\|^2 - \frac{\langle x,y\rangle^2}{\|y\|^2}\right)^{1/2}.$$

b) <u>Si $\langle x,y\rangle \leq 0$</u>, alors la projection de x sur $\langle y\rangle^+$ est 0, donc $d(x,\langle y\rangle^+) = \|x\|$.

<u>Si $\langle x,y\rangle > 0$</u>, alors la projection de x sur $\langle y\rangle^+$ est celle de x sur $\langle y\rangle$, donc d'après le rappel 2.2.1 on a :

$d(x,\langle y\rangle^+) = d(x,\langle y\rangle)$ ■

<u>Lemme</u> 2.2.3 : Si x et y sont deux points de $\mathbb{R}^m$, alors :

a) $\|y\|\cdot d(x,\langle y\rangle) = \|x\|\cdot d(y,\langle x\rangle)$

b) $\|y\|\cdot d(x,\langle y\rangle^+) = \|x\|\cdot d(y,\langle x\rangle^+)$.

<u>Démonstration</u> : a) cette égalité résulte directement

du point a) du lemme 2.2.2

b) - Si $\langle x,y\rangle \leqslant 0$, alors : $d(x,\langle y\rangle^{+}) = \|x\|$ et $d(y,\langle x\rangle^{+}) = \|y\|$,

donc : $\|y\|.d(x,\langle y\rangle^{+}) = \|y\|.\|x\| = \|x\|.d(y,\langle x\rangle^{+})$

- Si $\langle x,y\rangle > 0$, alors : $d(x,\langle y\rangle^{+}) = d(x,\langle y\rangle)$ et $d(y,\langle x\rangle^{+}) = d(y,\langle x\rangle)$,

donc : $\|y\|.d(x,\langle y\rangle^{+}) = \|y\|.d(x,\langle y\rangle) = \|x\|.d(y,\langle x\rangle) = \|x\|.d(y,\langle x\rangle^{+})$ ∎

Lemme 2.2.4 : $\forall\, x \in \mathbb{R}^m,\ \forall\, y \in \mathbb{R}^m,\ \forall\, z \in \mathbb{R}^m,$

$$d(x,\langle z\rangle^{+}).\,d(y,\langle x\rangle + \langle z\rangle) \leqslant d(y,\langle z\rangle^{+}).\,\|x-y\|.$$

Démonstration : Soit $\lambda \geqslant 0$. On a les inégalités suivantes:

(1) : $d(x,\langle z\rangle^{+}) \leqslant \|x-\lambda z\|$

(2) : $d(x-\lambda z,\langle y-\lambda z\rangle) \leqslant \|(x-\lambda z)-(y-\lambda z)\| = \|x-y\|.$

D'après le lemme 2.2.3 on a aussi

(3) : $\|x-\lambda z\|.\,d(y-\lambda z,\langle x-\lambda z\rangle) = \|y-\lambda z\|.\,d(x-\lambda z,\langle y-\lambda z\rangle)$,

donc en regroupant (1), (2) et (3), il vient :

(4) : $d(x,\langle z\rangle^{+}).\,d(y-\lambda z,\langle x-\lambda z\rangle) \leqslant \|y-\lambda z\|.\,\|x-y\|$,

et comme :

$\bullet\ d(y-\lambda z, <x>+<z>) \leqslant d(y-\lambda z, <x-\lambda z>)$

car $<x-\lambda z>$ est une partie de $<x>+<z>$,

$\bullet\ d(y, <x>+<z>) = d(y-\lambda z, <x>+<z>)$

car λz appartient au sous_espace vectoriel $<x>+<z>$,

il vient de (4) :

$$d(x,<z>^+)\cdot d(y,<x>+<z>) \leqslant \|y-\lambda z\|\cdot\|x-y\|.$$

En passant donc à la borne inférieure par arapport à λ on obtient :

$$d(x,<z>^+)\cdot d(y,<x>+<z>) \leqslant \inf_{\lambda\geqslant 0}\|y-\lambda z\|\cdot\|x-y\| = d(y,<z>^+)\cdot\|x-y\| \blacksquare$$

<u>Lemme</u> 2.2.5 : $\forall x\in\mathbb{R}^m$, $\forall y\in\mathbb{R}^m$, $\forall z\in\mathbb{R}^m$,

(5): $d(x,<z>^+)\cdot d(y,<x>^+ + <z>^+) \leqslant d(y,<z>^+)\cdot\|x-y\|$.

<u>Démonstration</u>: Si $x=0$ ou $y=0$, l'inégalité (5) est évidente. Si $z=0$, (5) résulte du lemme 2.2.3. Donc on peut supposer que x, y et z sont tous non nuls.

<u>1er cas: Si $<x,y>\leqslant 0$ ou $<y,z>\leqslant 0$.</u>

- Si $\langle x, y\rangle \leqslant 0$, alors $d(x,\langle y\rangle^+) = \|x\|$ et on a :

$$d(x,\langle z\rangle^+)\cdot d(y,\langle x\rangle^+ + \langle z\rangle^+) \leqslant \|x\|\cdot d(y,\langle z\rangle^+) = d(y,\langle z\rangle^+)\cdot d(x,\langle y\rangle^+)$$
$$\leqslant d(y,\langle z\rangle^+)\cdot \|x-y\|.$$

- Si $\langle y, z\rangle \leqslant 0$, alors $d(y,\langle z\rangle^+) = \|y\|$ et on a :

$$d(x,\langle z\rangle^+)\cdot d(y,\langle x\rangle^+ + \langle z\rangle^+) \leqslant \|x\|\cdot d(y,\langle x\rangle^+) = \|y\|\cdot d(x,\langle y\rangle^+)$$
$$= d(y,\langle z\rangle^+)\cdot d(x,\langle y\rangle^+)$$
$$\leqslant d(y,\langle z\rangle^+)\cdot \|x-y\|.$$

2eme Cas : Si $\langle x, y\rangle > 0$ et $\langle y, z\rangle > 0$.

a) Si x et z sont liés :

a.1) Si $x = \lambda z$ avec $\lambda > 0$ alors $d(x,\langle z\rangle^+) = 0$ et l'inégalité (5) est évidente.

a.2) Si $x = \lambda z$ avec $\lambda < 0$ alors $\langle x\rangle^+ + \langle z\rangle^+ = \langle x\rangle + \langle z\rangle$ et l'inégalité (5) résulte du lemme 2.2.4.

b) Si x et z ne sont pas liés :

Soit $\mathrm{vect}(x,z)$ le sous-espace vectoriel de $\mathbb{R}^m$ engendré par x et z. Il existe α, β appartenant à $\mathbb{R}$ et y' un élément de l'orthogonal de $\mathrm{vect}(x,z)$ tels que : $y = \alpha x + \beta z + y'$. Il vient donc :

$$(S): \begin{cases} \langle x,y\rangle = \alpha\,\|x\|^2 + \beta\,\langle x,z\rangle \\ \\ \langle z,y\rangle = \alpha\,.\,\langle x,z\rangle + \beta\,\|z\|^2 \end{cases}$$

Le déterminant du système (S) qui est égal à $\|x\|^2.\|z\|^2 - \langle x,z\rangle^2$ est non nul car x et z ne sont pas liés, donc :

$$\alpha = \frac{\|z\|^2.\langle x,y\rangle - \langle x,z\rangle.\langle y,z\rangle}{\|x\|^2.\|z\|^2 - \langle x,z\rangle^2} \quad \text{et} \quad \beta = \frac{\|x\|^2.\langle z,y\rangle - \langle x,z\rangle.\langle y,x\rangle}{\|x\|^2.\|z\|^2 - \langle x,z\rangle^2}$$

b.1) <u>si $\alpha \leqslant 0$ ou $\beta \leqslant 0$</u>

– <u>Si $\alpha \leqslant 0$</u>, la condition : " $\langle x,y\rangle > 0$ et $\langle y,z\rangle > 0$ " entraîne que $\langle x,z\rangle$ est strictement positif et par conséquent : $d(x,\langle z\rangle^+) = d(x,\langle z\rangle)$.

On a $d(y,\langle x\rangle^+ + \langle z\rangle^+) \leqslant d(y,\langle z\rangle^+)$, donc pour établir (5), il suffit de montrer que $d(x,\langle z\rangle) \leqslant \|x-y\|$.

On a : $\alpha \leqslant 0$, donc :

$$\|z\|^2.\langle x,y\rangle \leqslant \langle x,z\rangle.\langle y,z\rangle \leqslant \langle x,z\rangle.\|y\|.\|z\|,$$

Soit :

$$\|x\|^2 - \frac{\langle x,z\rangle^2}{\|z\|^2} \leqslant \|x\|^2 - \frac{\langle x,y\rangle^2}{\|y\|^2}.$$

D'où :

$$d(x,\langle z\rangle) \leqslant d(x,\langle y\rangle) \leqslant \|x-y\|.$$

– Si $\beta \leq 0$, dans ce cas aussi, la condition "$\langle x,y\rangle > 0$ et $\langle y,z\rangle > 0$" entraîne que $\langle x,z\rangle > 0$ et par conséquent $d(x,\langle z\rangle^+) = d(x,\langle z\rangle)$.

On a $d(y,\langle x\rangle^+ + \langle z\rangle^+) \leq d(y,\langle x\rangle^+) = d(y,\langle x\rangle)$, donc pour établir (5) il suffit de montrer que :

$$d(x,\langle z\rangle) \cdot d(y,\langle x\rangle) \leq d(y,\langle z\rangle) \cdot \|x-y\| .$$

On a $\beta \leq 0$, donc :

$$\|x\|^2 \cdot \langle y,z\rangle \leq \langle x,z\rangle \cdot \langle x,y\rangle \leq \langle x,z\rangle \cdot \|x\| \cdot \|y\| ,$$

Soit :

$$\|z\|^2 - \frac{\langle x,z\rangle^2}{\|x\|^2} \leq \|z\|^2 - \frac{\langle y,z\rangle^2}{\|y\|^2} .$$

D'où :

$$d(z,\langle x\rangle) \leq d(z,\langle y\rangle) .$$

En multipliant des deux cotés par $\dfrac{\|x\| \cdot d(y,\langle x\rangle)}{\|z\|}$, il vient enfin d'après le lemme 2.2.3 :

$$d(x,\langle z\rangle) \cdot d(y,\langle x\rangle) \leq d(y,z) \cdot d(x,\langle y\rangle) \leq d(y,\langle z\rangle) \cdot \|x-y\| .$$

b.2) si $\alpha > 0$ et $\beta > 0$.

Dans ce cas la projection de y sur $\langle x\rangle + \langle z\rangle$ (qui est égale à $\alpha x + \beta z$) appartient à $\langle x\rangle^+ + \langle z\rangle^+$, donc d'après le rappel 2.2.1, $d(y,\langle x\rangle^+ + \langle z\rangle^+) = d(y,\langle x\rangle + \langle z\rangle)$, et l'inégalité (5) résulte du lemme 2.2.4 ∎

<u>Preuve du lemme de réciprocité</u> : Supposons qu'il existe $N \geqslant 1$ tel que :

(6) : $\forall y \in H, d(y, H \cap K) \leqslant N \cdot d(y, K)$,

et montrons que :

(7) : $\forall x \in K, d(x, H \cap K) \leqslant N \cdot d(x, H)$.

Soit x un point de K. H est fermé donc il existe un élément y' de H tel que $d(x, H) = \|x - y'\|$.

<u>1er Cas</u> : <u>Si $y' \in K$</u> : Dans ce cas y' est un point de $K \cap H$, donc d'après le rappel 2.2.1, $d(x, H) = d(x, H \cap K)$, et comme $N \geqslant 1$, on peut écrire : $d(x, H \cap K) \leqslant N \cdot d(x, H)$.

<u>2ème cas</u> : <u>Si $y' \notin K$</u> : K étant un cône convexe, pour tout z de $H \cap K$, $\langle z \rangle^+$ est inclus dans $K \cap H$ et $\langle x \rangle^+ + \langle z \rangle^+$ est inclus dans K. Donc d'après le rappel 2.2.1 on a :

(8) : $\forall z \in H \cap K, d(x, H \cap K) \leqslant d(x, \langle z \rangle^+)$

(9) : $\forall z \in H \cap K, d(y', K) \leqslant d(y', \langle x \rangle^+ + \langle z \rangle^+)$.

Par ailleurs, d'après le lemme 2.2.5 on a :

(10) : $\forall z \in H \cap K, d(x, \langle z \rangle^+) \cdot d(y', \langle x \rangle^+ + \langle z \rangle^+) \leqslant d(y', \langle z \rangle^+) \cdot \|x - y'\|$.

En regroupant (8), (9) et (10), il vient donc :

$$\forall z \in H \cap K, d(x, H \cap K) \cdot d(y', K) \leqslant d(y', \langle z \rangle^+) \cdot \|x - y'\| = d(y', \langle z \rangle^+) \cdot d(x, H),$$

et si on passe à la borne inférieure par rapport à z :

$$d(x, H \cap K) \cdot d(y', K) \leqslant \inf_{z \in H \cap K} d(y', \langle z \rangle^+) \cdot d(x, H) = d(y', H \cap K) \cdot d(x, H).$$

D'après (6) il vient ensuite :

$$d(x, H \cap K) \cdot d(y', K) \leqslant N \cdot d(y', K) \cdot d(x, H),$$

et comme $d(y', K)$ est non nul car K est fermé et $y' \notin K$, il suffit de diviser par $d(y', K)$ pour obtenir (7) ■

L'existence de la constante N dans ce lemme n'est pas assureé pour des cônes quelconques :

<u>Exemple</u> : Dans $\mathbb{R}^3$, soient $H = \{(x, y, z) \in \mathbb{R}^3 / z = 0\}$ et $K = \mathbb{R}^+ \cdot S(a, 1)$ où $S(a, 1)$ est la sphère de centre $a = (1, 1, 1)$ et de rayon 1. H est un sous-espace vectoriel de $\mathbb{R}^3$ (c'est donc un cône convexe fermé) et K est un cône convexe fermé. On a $H \cap K = \langle e_2 \rangle^+$ où $e_2 = (0, 1, 0)$.

Supposons qu il existe une constante $N \geqslant 1$ telle que:

$$\forall\, x \in K,\ d(x, H\cap K) \leq N \cdot d(x, H).$$

Soit alors :

$$\forall\, (\alpha, \beta, \gamma) \in K,\ \sqrt{\alpha^2 + \gamma^2} \leq N \cdot |\gamma|.$$

Or on peut vérifier aisement que pour tout $\lambda > 0$, $(\lambda^2, \frac{\lambda + \lambda^3}{2}, \lambda)$ est un élément de K, donc pour tout $\lambda > 0$ on a $\lambda^2 \leq N^2 - 1$, ce qui est impossible.

Nous établissons l'existence de la constante N dans deux cas particuliers : quand les deux cônes sont deux sous-espaces vectoriels de $\mathbb{R}^m$ (proposition 2.2.6) et quand l'un des cônes est un sous-espace vectoriel de $\mathbb{R}^m$ et l'autre est un cône polyèdral de type $\mathbb{R}^{m+}_J$ (proposition 2.2.8).

<u>Proposition</u> 2.2.6 : Si H et K sont deux sous-espaces vectoriels de $\mathbb{R}^m$, alors il existe une constante $N \geq 1$ telle que :

$$\forall\, x \in K,\ d(x, K\cap H) \leq N \cdot d(x, H).$$

<u>Démonstration</u> : Dans l'espace quotient $K/_{K\cap H}$, les deux applications :

$\cdot\, N_1(\bar{x}) = d(x, K \cap H)$

$\cdot\, N_2(\bar{x}) = d(x, H)$

Sont deux normes, et comme la dimension de $K/_{K\cap H}$ est finie, il existe une constante $N > 0$ qu'on peut choisir supérieure à 1 telle que :

$$\forall x \in K \quad , \quad N_1(\bar{x}) \leq N \cdot N_2(\bar{x}) \, .$$

Soit :

$$\forall x \in K, \quad d(x, K \cap H) \leq N \cdot d(x, H) \quad \blacksquare$$

<u>Corollaire</u> 2.2.7 : Si H est un sous-espace vectoriel de $\mathbb{R}^m$ et I une partie de $[m]$, alors il existe une constante qu'on notera $N_I(H)$, qui soit la plus petite constante superieure à 1 verifiant :

$$\forall x \in \mathbb{R}^m_I \, , \quad d(x, \mathbb{R}^m_I \cap H) \leq N_I(H) \, . \, d(x, H) \, .$$

<u>Proposition</u> 2.2.8 : Si H est un sous-espace vectoriel de $\mathbb{R}^m$, alors pour toute partie J de $[m]$ on a :

$$(P^m_J) : \quad \forall y \in H \, , \, d(y, \mathbb{R}^{m+}_J \cap H) \leq \left[\sup_{J \subset I \subset [m]} N_I(H) \right]^{m-|J|} . \, d(y, \mathbb{R}^{m+}_J)$$

$$(q^m_J) : \quad \forall x \in \mathbb{R}^{m+}_J \, , \, d(x, \mathbb{R}^{m+}_J \cap H) \leq \left[\sup_{J \subset I \subset [m]} N_I(H) \right]^{m-|J|} . \, d(x, H) \, .$$

<u>Démonstration</u> : (P_J^m) et (q_J^m) étant équivalente par le lemme de réciprocité, il suffit de montrer l'une d'elles. La démonstration se fait par recurrence sur m. Pour tout entier positif m, on considère l'assertion suivante :

(a_m) : "Si H est un sous-espace vectoriel de $\mathbb{R}^m$, alors pour toute partie J de $[m]$, on a (P_J^m)".

• <u>(a_1) est vraie</u>, car les sous-espaces vectoriels de $\mathbb{R}$ sont $\{0\}$ et $\mathbb{R}$, et dans tous les cas, (P_J^1) est évidente.

• <u>Supposons que pour tout $n<m$, (a_n) est vraie</u>. Soient H un sous-espace vectoriel de $\mathbb{R}^m$ et J une partie de $[m]$. On fixe un élément y de H :

a) <u>Si $|J|=m$</u> (i.e $J=[m]$) alors $\mathbb{R}_J^{m+}=\{0\}$ et (P_J^m) est évidente.

b) <u>si $|J|\leq m-1$</u>, on a deux cas possibles :

b.1) <u>Si y appartient à $\mathbb{R}_J^{m+}$</u>, dans ce cas $d(y, \mathbb{R}_J^{m+}\cap H)$ et $d(y, \mathbb{R}_J^{m+})$ sont nuls et (P_J^m) est vraie.

b.2) <u>Si y n'appartient pas à $\mathbb{R}_J^{m+}$</u>, la projection de de y sur $\mathbb{R}_J^{m+}$ appartient à la frontière de $\mathbb{R}_J^{m+}$, donc il existe une partie K de $[m]$ telle que :

$$J\subset K,\ |K|=|J|+1 \text{ et } d(y, \mathbb{R}_J^{m+}) = d(y, \mathbb{R}_K^{m+}).$$

On a $\dim \mathbb{R}^m_K = m - |K| = m - |J| - 1 < m$ et $H_K = \mathbb{R}^m_K \cap H$ est un sous-espace vectoriel de $\mathbb{R}^m_K$, donc si on applique l'hypothèse de recurrence – en identifiant $\mathbb{R}^m_K$ à $\mathbb{R}^{m-|K|}$ – au sous-espace vectoriel H_K et à la partie vide de $[m]-K$, il vient :

$$(q_\phi^{m-|K|}) : \forall x \in \mathbb{R}^{m+}_K, \; d(x, \mathbb{R}^{m+}_K \cap H_K) \leqslant \Big[\sup_{\phi \subset L \subset [m]-K} N_L(H_K)\Big]^{m-|K|} \cdot d(x, H_K),$$

où pour toute partie L de $[m]-K$, $N_L(H_K)$ est la plus petite constante supérieure à 1 telle que :

$$\forall x \in \mathbb{R}^m_{K\cup L}, \; d(x, \mathbb{R}^m_{K\cup L} \cap H_K) \leqslant N_L(H_K) \cdot d(x, H_K).$$

<u>Lemme</u> : $\forall L \subset [m]-K$, $N_L(H_K) \leqslant N_{K\cup L}(H)$.

En effet, pour toute partie L de $[m]-K$, $N_{K\cup L}(H)$ vérifie :

$$\forall x \in \mathbb{R}^m_{K\cup L}, \; d(x, \mathbb{R}^m_{K\cup L} \cap H) \leqslant N_{K\cup L}(H) \cdot d(x, H),$$

et comme : $\mathbb{R}^m_{K\cup L} \cap H = \mathbb{R}^m_{K\cup L} \cap H_K$, il vient :

$$\forall x \in \mathbb{R}^m_{K\cup L}, \; d(x, \mathbb{R}^m_{K\cup L} \cap H_K) \leqslant N_{K\cup L}(H) \cdot d(x, H) \leqslant N_{K\cup L}(H) \cdot d(x, H_K).$$

Ce qui implique par définition de $N_L(H_K)$ (voir au dessus) que $N_L(H_K) \leqslant N_{K\cup L}(H)$.

Il vient donc de $(q_{\emptyset}^{m-|K|})$, d'après ce lemme :

$$\forall x \in \mathbb{R}_K^{m+},\ d(x, \mathbb{R}_K^{m+} \cap H) = d(x, \mathbb{R}_K^{m+} \cap H_K) \leqslant \Big[\sup_{\emptyset \subset L \subset [m]-K} N_{K \cup L}(H)\Big]^{m-|J|-1} . d(x, H_K)$$

$$\leqslant \Big[\sup_{K \subset I \subset [m]} N_I(H)\Big]^{m-|J|-1} . d(x, H_K)$$

$$\leqslant \Big[\sup_{J \subset I \subset [m]} N_I(H)\Big]^{m-|J|-1} . d(x, H_K),$$

et comme on a

$$\forall x \in \mathbb{R}_K^{m},\ d(x, H_K) = d(x, \mathbb{R}_K^{m} \cap H) \leqslant N_K(H) . d(x, H),$$

il vient en suite :

$$\forall x \in \mathbb{R}_K^{m+},\ d(x, \mathbb{R}_K^{m+} \cap H) \leqslant N_K(H) . \Big[\sup_{J \subset I \subset [m]} N_I(H)\Big]^{m-|J|-1} . d(x, H),$$

$$\leqslant \Big[\sup_{J \subset I \subset [m]} N_I(H)\Big]^{m-|J|} . d(x, H),$$

et par réciprocité :

$$\forall y' \in H,\ d(y', \mathbb{R}_K^{m+} \cap H) \leqslant \Big[\sup_{J \subset I \subset [m]} N_I(H)\Big]^{m-|J|} . d(y', \mathbb{R}_K^{m+}).$$

$\mathbb{R}^{m+}_K \cap H$ est une partie de $\mathbb{R}^{m+}_J \cap H$, donc :

$$\forall y' \in H,\ d(y', \mathbb{R}^{m+}_J \cap H) \leqslant d(y', \mathbb{R}^{m+}_K \cap H) \leqslant \Big[\sup_{J \subset I \subset [m]} N_I(H)\Big]^{m-|J|} . d(y', \mathbb{R}^{m+}_K),$$

en particulier pour $y' = y$, on a :

$$d(y, \mathbb{R}^{m+}_J \cap H) \leqslant \Big[\sup_{J \subset I \subset [m]} N_I(H)\Big]^{m-|J|} . d(y, \mathbb{R}^{m+}_K) = \Big[\sup_{J \subset I \subset [m]} N_I(H)\Big]^{m-|J|} . d(y, \mathbb{R}^{m+}_J).$$

D'où (P^m_J) ■

2.3 Démonstration du théorème d'approximation.

Rappels et notations : Si V est une application linéaire continue d'un Banach E dans $\mathbb{R}^m$, on note $\bar{V}$ l'isomorphisme induit par V de $E/_{\mathrm{Ker}V}$ dans $\mathrm{Im}\,V$ et $\mu(V) = \| \bar{V}^{-1} \|$.

$$\begin{tikzcd} E \arrow[r, "V"] \arrow[d, "p"'] & \mathbb{R}^m \\ E/_{\mathrm{Ker}V} \arrow[r, "\bar{V}"'] & \mathrm{Im}\,V \arrow[u, hook, "i"'] \end{tikzcd}$$

On a alors :

- $\forall x \in E,\ \|V(x)\| \leqslant \|V\| . d(x, \mathrm{Ker}\,V)$
- $\forall x \in E,\ \|p(x)\|_{E/_{\mathrm{Ker}V}} = d(x, \mathrm{Ker}V) \leqslant \mu(V) . \|V(x)\|.$

On va être amené à utiliser la proposition 2.2.8 dans le cas où H est l'image de V. On a alors :

<u>Proposition</u> 2.3.1 : Soit V une application linéaire continue d'un Banach E dans $\mathbb{R}^m$. Alors pour toute partie J de $[m]$ on a :

$$\forall x \in E,\ d(x, V^{-1}(\mathbb{R}^{m+}_J)) \leq \mu(V) \cdot \left[\sup_{J \subset I \subset [m]} N_I(\operatorname{Im} V)\right]^{m-|J|} \cdot d(V(x), \mathbb{R}^{m+}_J).$$

<u>Démonstration</u> : Soit J une partie de $[m]$, d'après la proposition 2.2.8 on a :

$$\forall y \in \operatorname{Im} V,\ d(y, \mathbb{R}^{m+}_J \cap \operatorname{Im} V) \leq \left[\sup_{J \subset I \subset [m]} N_I(\operatorname{Im} V)\right]^{m-|J|} \cdot d(y, \mathbb{R}^{m+}_J) .$$

Soit :

$$\forall x \in E,\ d(V(x), \mathbb{R}^{m+}_J \cap \operatorname{Im} V) \leq \left[\sup_{J \subset I \subset [m]} N_I(\operatorname{Im} V)\right]^{m-|J|} \cdot d(V(x), \mathbb{R}^{m+}_J).$$

<u>Lemme</u> : Pour toute partie C non vide de $\mathbb{R}^m$ et pour tout élément x de E on a : $d(x, V^{-1}(C)) \leq \mu(V) \cdot d(V(x), \operatorname{Im} V \cap C)$.

En effet on a :

$$d(x, V^{-1}(C)) = \inf_{V(x')\in C} \|x - x'\| = \inf_{V(x')\in C} \|p(x) - p(x')\|$$

$$\leqslant \inf_{V(x')\in C} \|\bar{V}^{-1}\| \cdot \|\bar{V}(p(x)) - \bar{V}(p(x'))\|$$

$$= \mu(V) \cdot \inf_{V(x')\in C} \|V(x) - V(x')\|$$

$$= \mu(V) \inf_{y \in \mathrm{Im}V \cap C} \|V(x) - y\|$$

$$= \mu(V) \cdot d(V(x), \mathrm{Im}V \cap C) .$$

En prenant donc $C = \mathbb{R}^{m+}_J$, il vient :

$$\forall x \in E, d(x, V^{-1}(\mathbb{R}^{m+}_J)) \leqslant \mu(V) \cdot d(V(x), \mathrm{Im}V \cap \mathbb{R}^{m+}_J)$$

$$\leqslant \mu(V) \cdot \left[\sup_{J \subset I \subset [m]} N_I(\mathrm{Im}V)\right]^{m-|J|} \cdot d(V(x), \mathbb{R}^{m+}_J) \blacksquare$$

2.3.2 <u>Preuve du théorème d'approximation</u> : Les notations sont celles du théorème.

1) L'existence des constantes N_I est établie par le Corollaire 2.2.7. Pour tout $I \subset [m]$ on prend $N_I = N_I(\mathrm{Im}Dh(a))$.

2) Soit $\varepsilon > 0$ et x un élément de $W^* \cap \bar{B}(a, \frac{2\varepsilon}{M.A})$. La formule de Taylor à l'ordre 1 pour h au point a donne :

$$h(x)-Dh(a)\cdot(x-a)=\int_0^1 [Dh(a+t(x-a))-Dh(a)]\cdot(x-a)\cdot dt.$$

D'où :

$$\|h(x)-Dh(a)\cdot(x-a)\|\leqslant\int_0^1 \|Dh(a+t(x-a))-Dh(a)\|\cdot\|x-a\|\cdot dt.$$

Ω étant étoilé au point a, pour tout t de $]0,1[$, $a+t(x-a)$ appartient à $\Omega-\{a\}$, donc avec la condition de croissance que vérifie Dh au point a, il vient :

$$(1) : \|h(x)-Dh(a)\cdot(x-a)\|<\int_0^1 A.t.\|x-a\|.\|x-a\|.dt=\frac{A}{2}\|x-a\|^2.$$

D'autres parts, d'après la proposition 2.3.1 on a :

$$\forall x'\in E,\ d(x',Q_a)=d(x',Dh(a)^{-1}(\mathbb{R}^{m+}_{[p]}))\leqslant M.d(Dh(a).x',\mathbb{R}^{m+}_{[p]}),$$

en particulier pour $x'=x-a$:

$$d(x-a,Q_a)\leqslant M\cdot d(Dh(a)\cdot(x-a),\mathbb{R}^{m+}_{[p]}),$$

et comme $h(x)$ appartient à $\mathbb{R}^{m+}_{[p]}$, il vient en suite :

$$(2) : d(x-a,Q_a)\leqslant M\cdot\|Dh(a)\cdot(x-a)-h(x)\|.$$

En regroupant (1) et (2) on obtient alors : $d(x-a,Q_a)<\frac{M\cdot A}{2}\|x-a\|^2$, et comme $\|x-a\|\leqslant\frac{2\varepsilon}{M\cdot A}$, il résulte enfin : $d(x-a,Q_a)<\varepsilon.\|x-a\|$, c'est-à-dire que x appartient à $a+(Q_a)_\varepsilon$ ∎

2.4 Évaluation de la constante M.

Les notations sont celles du théorème d'approximation. Le calcul exact de M s'avère difficile même dans le cas où la dimension de E est finie. On est amené alors à remplacer M par un majorant M' qui est calculable et qui diffère aussi peu que possible de M. Ce faisant on perd un peu de précision sur la localisation de W au voisinage de a, mais on détermine explicitement une application $\varepsilon \longmapsto r(\varepsilon) = \frac{2 \cdot \varepsilon}{M' \cdot A}$ qui permet la mesure de l'approximation.

Rappelons que $M = \mu(Dh(a)) \cdot \left[\sup_{[p] \subset I \subset [m]} N_I(\operatorname{Im} Dh(a))\right]^{|I_0|}$. Nous allons établir en général des majorations explicites de $\mu(V)$ et de $N_I(\operatorname{Im} V)$, où V est une application linéaire continue d'un Banach E dans $\mathbb{R}^m$ et I une partie de $[m]$, une majoration calculable de M en découlera.

2.4.1 Évaluation de $\mu(V)$.

Pour évaluer $\mu(V)$ en dimension finie, on a le résultat suivant:

Proposition 2.4.1.1 : Soient V une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$, ${}^tV$ sa transposée et $Sp({}^tV.V)$ l'ensemble des valeurs propres de ${}^tV.V$. Alors :

1) $\|V\| = \left[\sup_{\lambda \in Sp({}^tV.V)} \lambda\right]^{1/2}$

2) $\mu(V) = \left[\sup_{\substack{\lambda \in Sp({}^tV.V) \\ \lambda \neq 0}} \frac{1}{\lambda}\right]^{1/2}$

Démonstration : 1) voir [6] par exemple.

2) Soit $(e_1, e_2, \dots e_d, e_{d+1}, \dots e_n)$ une base orthogonale de $\mathbb{R}^n$ telle que $(e_{d+1}, \dots e_n)$ est une base de $\mathrm{Ker}\, V$ et soit $(f_1, f_2, \dots f_d, f_{d+1}, \dots f_m)$ une base de $\mathbb{R}^m$ telle que $(f_1, f_2, \dots f_d)$ est une une base de $\mathrm{Im}\, V$. Les matrices de V et de ${}^tV.V$ dans ces deux bases s'écrivent respectivement:

$$\left[\begin{array}{c|c} U & 0 \\ \hline 0 & 0 \end{array}\right] \quad \text{et} \quad \left[\begin{array}{c|c} {}^tU.U & 0 \\ \hline 0 & 0 \end{array}\right]$$

où U est la matrice dans les deux bases $(e_1, e_2, \dots e_d)$ et $(f_1, f_2, \dots f_d)$ de la restriction de V sur $(\mathrm{Ker}\, V)^\perp$.

Montrons d'abord que $\mu(V) = \mu(U)$.

Pour tout élément x de E il existe un élément x_1 de $\ker V$ et un élément x_2 de $(\ker V)^\perp$ tels que $x = x_1 + x_2$, et on a :

$$d(x, \ker V) = \|x_2\| \leq \mu(U) \cdot \|U(x_2)\| = \mu(U) \cdot \|V(x)\| ,$$

et comme $\mu(v)$ est la plus petite constante qui vérifie :

$$\forall x \in E, \quad d(x, \ker v) \leqslant \mu(v) \cdot \|v(x)\|,$$

il résulte que $\mu(v) \leqslant \mu(u)$.

Pour tout élément y de $(\ker v)^{\perp}$, on a :

$$d(y, \ker u) = \|y\| = d(y, \ker v) \leqslant \mu(v) \cdot \|v(y)\| = \mu(v) \cdot \|u(y)\|,$$

et comme $\mu(u)$ est la plus petite constante qui vérifie :

$$\forall y \in (\ker v)^{\perp}, d(y, \ker u) \leqslant \mu(u) \cdot \|u(y)\|,$$

il résulte que $\mu(u) \leqslant \mu(v)$. Donc $\mu(u) = \mu(v)$.

u est bijective donc :

$$\mu(u) = \|u^{-1}\| = \left[\sup_{\lambda \in Sp(u^{-1} \cdot {}^t u^{-1})} \lambda \right]^{1/2},$$

et comme :

$$\cdot \ Sp(u^{-1} \cdot {}^t u^{-1}) = \{ 1/\lambda \ / \ \lambda \in Sp({}^t u \cdot u) \}$$

$$\cdot \ Sp({}^t u \cdot u) = Sp({}^t v \cdot v) - \{0\}$$

il résulte :

$$\mu(v) = \mu(u) = \left[\sup_{\substack{\lambda \in Sp({}^t v \cdot v) \\ \lambda \neq 0}} 1/\lambda \right]^{1/2} \blacksquare$$

${}^t v.v$ est représenté dans une base de $\mathbb{R}^n$ par une matrice (a_{ij}) et les valeurs propres de ${}^t v.v$ sont obtenues comme racines du polynôme caractéristique de degré n: $P(x) = \det(x.\mathrm{Id} - a_{ij})$. On ne sait pas toujours calculer ces racines mais on peut majorer $\|v\|$ et $\mu(v)$ si on connaît le coefficient de x^{n-1} et les deux derniers coefficients non nuls de $P(x)$:

<u>Proposition</u> 2.4.1.2 : Soient v une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ et $P(x) = x^n - a_1 x^{n-1} + \dots + (-1)^{p-1} . a_{p-1} x^{n-p+1} + (-1)^p a_p x^{n-p}$, le polynôme caractéristique de ${}^t v.v$ où a_p est le dernier coefficient non nul de $P(x)$. Alors :

1) $\|v\|^2 \leqslant a_1$

2) $\mu(v)^2 \leqslant \dfrac{a_{p-1}}{a_p}$

<u>Demonstration</u> : 1) Les valeurs propres $\lambda_1, \lambda_2, \dots \lambda_k$ de ${}^t v.v$ sont toutes réelles et positives et elles sont racines du polynôme $P(x)$, donc :

$$a_1 = \lambda_1 + \lambda_2 + \dots + \lambda_k \geqslant \sup \lambda_i = \|v\|^2 .$$

2) On a :

$$P(x) = x^{n-p}(x^p - a_1 x^{p-1} + \dots + (-1)^{p-1} a_{p-1} x + (-1)^p a_p) .$$

Si on fait le changement de variable $Y = \frac{1}{X}$, les inverses des valeurs propres non nulles deviennent racines du polynôme $Q(Y) = 1 - a_1 Y + \dots + (-1)^{p-1} a_{p-1} Y^{p-1} + (-1)^p a_p Y^p$, donc

$$-\frac{(-1)^{p-1} a_{p-1}}{(-1)^p a_p} = \frac{a_{p-1}}{a_p} = \sum_{\lambda_i \neq 0} \frac{1}{\lambda_i} \geqslant \sup_{\lambda_i \neq 0} \frac{1}{\lambda_i} = \mu(V)^2 \blacksquare$$

Dans le cas où la dimension de E est infinie, on peut ramener le problème de majoration de $\mu(V)$ à un problème de calcul de valeurs propres en dimension finie, à condition de connaître explicitement une base normé d'un supplémentaire de ker V :

<u>Proposition</u> 2.4.1.3 : Soient V une application linéaire continue d'un Banach E dans $\mathbb{R}^m$, S un supplémentaire de Ker V (il existe au moins un car la codimension de Ker V est finie) et $(b_1, b_2, \dots b_n)$ une base normée de S. Soit $\tilde{V}$ l'application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ définie par :

$$(x_1, x_2, \dots x_n) \longmapsto \tilde{V}(x_1, x_2, \dots x_n) = \sum_{i=1}^{n} x_i V(b_i).$$

Alors :

1) $\mu(V) \leqslant \sqrt{m} \cdot \mu(\tilde{V})$

2) Si de plus E muni de sa norme est un Hilbert, si $S = (\ker V)^{\perp}$ et si $(b_1, b_2, \dots b_n)$ est orthogonale alors $\mu(V) = \mu(\tilde{V})$.

<u>Démonstration</u> : 1) Pour tout élément x de E, il existe un élément x_1 de $\ker V$, un élément x_2 de S tels que $x = x_1 + x_2$, et il existe $\alpha_1, \alpha_2, \dots \alpha_n$ appartenant à $\mathbb{R}$ tels que $x_2 = \sum_{i=1}^{n} \alpha_i b_i$. Si on note $\tilde{x}_2 = (\alpha_1, \alpha_2, \dots \alpha_n)$, on a :

$$(1) : \begin{cases} V(x_2) = \tilde{V}(\tilde{x}_2) \\ \|x_2\| = \|\sum_{i=1}^{n} \alpha_i b_i\| \leqslant \sum_{i=1}^{n} |\alpha_i| \leqslant \sqrt{m}\, \|\tilde{x}_2\|_2 \end{cases}$$

où $\|\,.\,\|_2$ est la norme euclididienne dans $\mathbb{R}^m$. Il vient donc :

$$d(x, \ker V) \leqslant \|x_2\| \leqslant \sqrt{m}\, \|\tilde{x}_2\|_2 \leqslant \sqrt{m}\, \mu(\tilde{V}) \|\tilde{V}(\tilde{x}_2)\| = \sqrt{m}\, \mu(\tilde{V})\, \|V(x_2)\|$$
$$= \sqrt{m}\, \mu(\tilde{V})\, \|V(x)\|,$$

et comme $\mu(V)$ est la plus petite constante qui vérifie :

$$\forall x \in E, \; d(x, \ker V) \leqslant \mu(V) \,.\, \|V(x)\|,$$

il résulte que $\mu(V) \leqslant \sqrt{m}\, \mu(\tilde{V})$.

2) Si E est un Hilbert, $S = (\ker V)^{\perp}$ et $(b_1, b_2, \dots b_n)$ est orthogonale alors en reprenant les notations précédentes, (1) devient :

$$\begin{cases} V(x_2) = \tilde{V}(\tilde{x}_2) \\ \|x_2\| = \|\sum_{i=1}^{n} \alpha_i b_i\| = \left(\sum_{i=1}^{n} \alpha_i^2\right)^{1/2} = \|\tilde{x}_2\|_2 . \end{cases}$$

donc :

$$\forall x \in E \;,\quad d(x, \ker V) \leqslant \psi(\tilde{V}) \cdot \|V(x)\| \;,$$

et il résulte que $\psi(V) \leqslant \psi(\tilde{V})$.

D'autres parts, pour tout $y = (y_1, y_2, \dots y_n)$ de $\mathbb{R}^n$, on a :

$$\|y\|_2 = d\Big(\sum_{i=1}^{n} y_i b_i, \ker V\Big) \leqslant \psi(V) \cdot \Big\|V\Big(\sum_{i=1}^{n} y_i b_i\Big)\Big\|$$
$$= \psi(V) \; \|\tilde{V}(y)\| \;,$$

et comme $\psi(\tilde{V})$ est la plus petite constante qui vérifie :

$$\forall y \in \mathbb{R}^n, \; \|y\| \leqslant \psi(\tilde{V}) \cdot \|\tilde{V}(y)\| \;,$$

il résulte que $\psi(\tilde{V}) \leqslant \psi(V)$. Donc $\psi(V) = \psi(\tilde{V})$ ∎

2.4.2 Évaluation de $N_I(\operatorname{Im} V)$

Pour majorer $N_I(\operatorname{Im} V)$, on va d'abord expliciter une constante N qui vérifie la proposition 2.2.6 en prenant $H = \operatorname{Im} V$ et $K = \ker p$, où p est une application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^k$:

<u>Lemme</u> 2.4.2.1 : Si V est une application linéaire continue du Banach E dans $\mathbb{R}^m$ et p une application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^k$, alors :

$$\forall x \in \operatorname{Ker} p,\ d(x, \operatorname{Ker} p \cap \operatorname{Im} V) \leqslant N_1 . d(x, \operatorname{Im} V),$$

où $N_1 = \|p\| . \|V\| . \mu(p \circ V)$.

<u>Démonstration</u> :

$$\begin{tikzcd}
E \arrow[r, "V"] \arrow[d, "\rho"'] & \mathbb{R}^m \arrow[r, "p"] \arrow[d, "\pi"'] & \mathbb{R}^k \\
E/_{\ker p\circ V} \arrow[r, "V'"] \arrow[rrd, "\overline{p\circ V}"'] & \mathbb{R}^m/_{\ker p} \arrow[r, "\bar{p}"] & \operatorname{Im} p \arrow[u, hook] \\
 & & \operatorname{Im} p\circ V \arrow[u, hook]
\end{tikzcd}$$

Notons d'abord que si $x \equiv x' \pmod{\ker p\circ V}$ alors $V(x) \equiv V(x') \pmod{\ker p}$. D'où les factorisations :

$V' \circ \rho = \pi \circ V$ et $\overline{p\circ V} = \bar{p} \circ V'$.

Pour tout élément x de E on a :

$$d(V(x), \operatorname{Ker} p \cap \operatorname{Im} V) = \inf_{p\circ V(x')=0} \|V(x) - V(x')\| \leqslant \|V\| . \inf_{p\circ V(x')=0} \|x - x'\|$$

$$= \|V\| . \|\rho(x)\|_{E/_{\ker p\circ V}}$$

$$\leqslant \|V\| . \mu(p\circ V) . \|\overline{p\circ V}(\rho(x))\|$$

$$
\begin{aligned}
d(v(x), \ker p \cap \mathrm{Im} v) &\leq \|v\| \cdot \gamma(p\circ v) \cdot \|\overline{p\circ v}(\Delta(x))\| \\
&= \|v\| \cdot \gamma(p\circ v) \cdot \|\bar{p}\circ v'(\Delta(x))\| \\
&\leq \|v\| \cdot \gamma(p\circ v) \cdot \|\bar{p}\| \, \|v'(\Delta(x))\|_{\mathbb{R}^m/\ker p} \\
&= \|v\| \cdot \gamma(p\circ v) \cdot \|p\| \cdot d(v(x), \ker p).
\end{aligned}
$$

Soit :

$\forall y \in \mathrm{Im} v$, $d(y, \ker p \cap \mathrm{Im} v) \leq \|v\| \cdot \|p\| \cdot \gamma(p\circ v) \cdot d(y, \ker p)$,

et d'après le lemme de réciprocité, il vient :

$\forall x \in \ker p$, $d(x, \ker p \cap \mathrm{Im} v) \leq \|v\| \cdot \|p\| \cdot \gamma(p\circ v) \cdot d(x, \mathrm{Im} v)$ ∎

<u>Remarque</u> 2.4.2.2 : Soit $\tilde{v}$ l'application linéaire définie dans la proposition 2.4.1.3. Alors on a aussi :

$\forall x \in \ker p$, $d(x, \ker p \cap \mathrm{Im} v) \leq N_2 \, d(x, \mathrm{Im} v)$,
où $N_2 = \|p\| \cdot \|\tilde{v}\| \cdot \gamma(p\circ\tilde{v})$.

<u>En effet</u>, d'après le lemme précédent on a :

$\forall x \in \ker p$, $d(x, \ker p \cap \mathrm{Im} \tilde{v}) \leq \|p\| \cdot \|\tilde{v}\| \cdot \gamma(p\circ\tilde{v}) \cdot d(x, \mathrm{Im} \tilde{v})$,

et comme $\mathrm{Im} v = \mathrm{Im} \tilde{v}$, on peut écrire :

$\forall x \in \ker p$, $d(x, \ker p \cap \mathrm{Im} v) \leq N_2 \, d(x, \mathrm{Im} v)$ ∎

Corollaire 2.4.2.3 : Les hypothèses sont celles de la proposition 2.4.1.3. Pour toute partie I de $[m]$ on a :

1) $N_I(\operatorname{Im} v) \leqslant \|v\| \cdot \mu(v_I)$

2) $N_I(\operatorname{Im} v) \leqslant \|\tilde{v}\| \cdot \mu(\tilde{v}_I)$

Démonstration : 1) Supposons que $I = \{i_1, i_2, \dots i_k\}$ avec $i_1 < i_2 < \dots < i_k$. Si on note p_I la projection linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^k$ définie par :

$$y = (y_1, y_2, \cdots y_m) \longmapsto p_I(y) = (y_{i_1}, y_{i_2}, \cdots, y_{i_k})$$

alors on a : $\operatorname{Ker} p_I = \mathbb{R}^m_I$ et $v_I = p_I \circ v$. Donc d'après le lemme 2.4.2.1 :

$$\forall x \in \mathbb{R}^m_I \, , \, d(x, \mathbb{R}^m_I \cap \operatorname{Im} v) \leqslant \|v\| . \|p_I\| \cdot \mu(p_I \circ v) . d(x, \operatorname{Im} v) .$$

Soit :

$$\forall x \in \mathbb{R}^m_I \, , \, d(x, \mathbb{R}^m_I \cap \operatorname{Im} v) \leqslant \|v\| . \mu(v_I) . d(x, \operatorname{Im} v) \, ,$$

ce qui implique d'après la définition de $N_I(\operatorname{Im} v)$ que :
$N_I(\operatorname{Im} v) \leqslant \|v\| \cdot \mu(v_I)$.

2) D'après la remarque 2.4.2.2 on a :

$$\forall x \in \mathbb{R}^m_I \,,\, d(x, \mathbb{R}^m_I \cap \mathrm{Im}\, v) \leqslant \|\tilde{v}\| \cdot \|P_I\| \cdot \gamma(P_I \circ \tilde{v}) \cdot d(x, \mathrm{Im}\, v).$$

Soit :

$$\forall x \in \mathbb{R}^m_I \,,\, d(x, \mathbb{R}^m_I \cap \mathrm{Im}\, v) \leqslant \|\tilde{v}\| \cdot \mu(\tilde{v}_I) \cdot d(x, \mathrm{Im}\, v) \,,$$

et d'après la définition de $N_I(\mathrm{Im}\, v)$ on a $N_I(\mathrm{Im} v) \leqslant \|\tilde{v}\| \cdot \mu(\tilde{v}_I)$.

2.4.3 Application à l'explicitation d'un majorant de M.

D'après le corollaire 2.4.2.3 on obtient donc la majoration suivante :

$$(1) : \; M \leqslant M' = \mu(Dh(a)) \cdot \Big[\|Dh(a)\| \cdot \sup_{[p] \subset I \subset [m]} \mu(Dh(a)_I) \Big]^{|I_0|},$$

qu'on peut évaluer directement par la proposition 2.4.1.1 (et éventuellement par la proposition 2.4.1.2) si la dimension de E est finie.

Si la dimension de E est infinie et si on connaît explicitement une base normée $(b_1, b_2, \ldots, b_n)$ d'un supplémentaire S de Ker Dh(a), alors en notant $\widetilde{Dh}(a)$ l'application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ définie par :

$$\widetilde{Dh}(a).(y_1, y_2, \dots, y_n) = \sum_{i=1}^{n} y_i \, Dh(a)(b_i),$$

on obtient d'après la proposition 2.4.1.3 et le Corollaire 2.4.2.3 les deux majorations suivantes (qu'on peut aussi évaluer par la proposition 2.4.1.1) :

$$(2) : M \leqslant \sqrt{m} \,.\, \mu(\widetilde{Dh}(a)) \,.\, \left[\| \widetilde{Dh}(a) \| \,.\, \sup_{[p] \subset I \subset [m]} \mu(\widetilde{Dh}(a)_I) \right]^{|I_0|},$$

et si E muni de sa norme est un Hilbert, $S = (\operatorname{Ker} Dh(a))^{\perp}$ et $(b_1, b_2, \dots, b_n)$ est orthogonale :

$$(3) : M \leqslant \mu(\widetilde{Dh}(a)) \,.\, \left[\| \widetilde{Dh}(a) \| \,.\, \sup_{[p] \subset I \subset [m]} \mu(\widetilde{Dh}(a)_I) \right]^{|I_0|}.$$

CHAPITRE. III

APPLICATION À L'ÉTUDE D'UN PROBLÈME EN OPTIMISATION PARAMÈTRÉ

Soit J une application numérique de classe C^3, définie sur un ouvert $V \times \Omega$ de $\mathbb{R}^n \times \mathbb{R}^p$. On considère le problème d'optimisation paramètré suivant :

$\mathcal{P}(\alpha)$: , minimiser $J(\alpha, \cdot)$ sur Ω .

x est un point critique pour $\mathcal{P}(\alpha)$ si x vérifie les deux conditions nécessaires suivantes :

(1) : $D_x J(\alpha, x) = 0$

(2) : $D^2_{xx} J(\alpha, x) \cdot (X, X) \geqslant 0 \quad , \forall X \in \mathbb{R}^p .$

<u>Lemme</u> : Soit $P(X) = \det(X \cdot Id - A) = X^p + \sum_{i=1}^{p} (-1)^i a_i X^{p-i}$, le polynôme caractéristique d'une matrice symétrique A de $\mathbb{R}^p$. Alors A est positive (c.à.d : $\langle A \cdot X, X \rangle \geqslant 0 \ \forall X \in \mathbb{R}^p$) ssi :

$\forall i \in [p] \quad , \ a_i \geqslant 0$

Démonstration : Soient $\alpha_1, \alpha_2, \ldots \alpha_p$ les valeurs propres de A . A est positive ssi $\alpha_1, \alpha_2, \dots \alpha_p$ sont toutes positives, donc il suffit de montrer l'équivalence suivante :

$$(i) : (\forall k \in [p] , \alpha_k \geqslant 0) \iff (ii) : (\forall k \in [p] , a_k \geqslant 0).$$

Soient $S_k(x_1, x_2, \dots, x_p) = \sum_{i_1 < i_2 < \dots < i_k} x_{i_1} . x_{i_2} \dots x_{i_k}$, $k = 1, 2, \dots p$ les p polynômes symétriques élémentaire à p variables. $\alpha_1, \alpha_2, \dots \alpha_p$ sont racines du polynôme $P(x)$, donc :

$$\forall k \in [p] , \quad a_k = S_k(\alpha_1, \alpha_2, \dots, \alpha_p).$$

ona évidement

$$(\forall k \in [p] , \alpha_k \geqslant 0) \Rightarrow (\forall k \in [p] , S_k(\alpha_1, \alpha_2, \dots \alpha_p) \geqslant 0),$$

donc (i) implique (ii).

Supposons maintenant que (ii) est vérifieé et que α_1 par exemple est strictement négative. on a :

$$P(\alpha_1) = \alpha_1^p + \sum_{i=1}^{p} (-1)^i a_i \, \alpha_1^{p-i} = 0$$

donc :

$$1 = \sum_{i=1}^{p} - a_i . \frac{1}{(-\alpha_1)^i} ,$$

ce qui est impossible car les termes de cette sommation sont tous négatifs. Donc (ii) implique (i).

Si pour tout (α, x) de $V \times \Omega$ on associe les coefficients $g_1(\alpha, x), g_2(\alpha, x), \dots g_p(\alpha, x)$ tels que :

$$P(X) = X^p + \sum_{i=1}^{p} (-1)^i g_i(\alpha, x) . X^{p-i}$$

soit le polynôme caractéristique de la matrice symétrique associée à la forme quadratique $D^2_{xx} J(\alpha, x)$, alors d'après le lemme précédent, l'inégalité (2) est équivalente à :

(3) : $\forall i \in [p]$, $g_i(\alpha, x) \geqslant 0$.

Soit W la partie de $V \times \Omega$ définie par :

$$W = \{ (\alpha, x) \in V \times \Omega \,/\, D_x J(\alpha, x) = 0 \text{ et } \forall i \in [p], g_i(\alpha, x) \geqslant 0 \} .$$

Alors pour tout α de V et tout x point critique pour $\mathcal{P}(\alpha)$, le couple (α, x) appartient à W. Réciproquement si (α, x) appartient à W alors x est un point critique pour $\mathcal{P}(\alpha)$.

Afin d'obtenir un contrôle sur l'ensemble des points critiques de $\mathcal{P}(\alpha)$ au voisinage d'un paramètre α_0 de V, nous allons appliquer le résultat du chapitre II à l'ensemble W au point $a = (\alpha_0, x_0)$ où x_0 est un point critique pour $\mathcal{P}(\alpha_0)$.

On pose : $f = \left(\frac{\partial J}{\partial x_i} \right)_{i \in [p]}$, $I_0 = \{ i \in [p] \,/\, g_i(\alpha_0, x_0) = 0 \}$,

$g_0 = (g_i)_{i \in I_0}$, $h = (f, g_0)$ et $m = p + |I_0|$.

L'application J est de classe C^3, donc f est de classe C^2, g_0 est de classe C^1 et par conséquent h est de classe C^1. On a le résultat suivant :

<u>Proposition</u> 3.1 : Soient $Q_a = \{X \in \mathbb{R}^n \times \mathbb{R}^p / Df(a).X = 0 \text{ et } \forall i \in I_0, Dg_i(a).X \geqslant 0\}$, le cône des conditions d'ordre 1 de W au point $a = (\alpha_0, x_0)$, et A une constante positive telle que :

$$\forall (\alpha, x) \in V \times \Omega,\ \|Dh(\alpha, x) - Dh(\alpha_0, x_0)\| \leqslant A \,\|(\alpha - \alpha_0, x - x_0)\|.$$

Alors pour tout α de V et tout x point critique pour $P(\alpha)$, on a :

$$d\big((\alpha - \alpha_0, x - x_0), Q_a\big) \leqslant \frac{M'.A}{2} \,.\, \|(\alpha - \alpha_0, x - x_0)\|^2,$$

où $M' = \mu(Dh(a)).\left[\|Dh(a)\|.\sup_{[p] \subset I \subset [m]} \mu(Dh(a)_I)\right]^{|I_0|}$

<u>Démonstration</u> : Si on approxime W par Q_a au point a, il vient d'après le théorème d'approximation et la majoration (1) du paragraphe 2.4.3 (cf ch II) :

$$\forall \varepsilon > 0,\ W^* \cap \bar{B}\left(a, \frac{2\varepsilon}{M'.A}\right) \subset a + (Q_a)_\varepsilon,$$

et d'après la remarque 1.1.2 du chapitre I, si on pose $e(\eta)=\frac{M'.A}{2}.\eta$, il vient :

$$\forall \eta > 0 \ , \ W^* \cap \bar{B}(a, \eta) \subset a + (Q_a)_{e(\eta)}$$

et par suite :

$$W^* \subset (\alpha_0, x_0) + \tilde{Q}_a$$

où $\tilde{Q}_a$ est le horn- neighborhood de Q_a défini par :

$$\tilde{Q}_a = \{(\alpha, x) \in \mathbb{R}^n \times \mathbb{R}^p / \ d((\alpha, x), Q_a) \leqslant e(\|(\alpha, x)\|) . \|(\alpha, x)\| \}$$

$$= \{(\alpha, x) \in \mathbb{R}^n \times \mathbb{R}^p / \ d((\alpha, x), Q_a) \leqslant \frac{M'.A}{2} \ \|(\alpha, x)\|^2 \} ,$$

et comme pour tout α de V et tout x point critique pour $f(\alpha)$, (α, x) appartient à W, il vient :

$$d((\alpha-\alpha_0, x-x_0), Q_a) \leqslant \frac{M'.A}{2} \ \|(\alpha-\alpha_0, x-x_0)\|^2 \ \blacksquare$$

<u>Exemple 1</u> : Soit J l'application numérique définie sur $\mathbb{R}^2 \times \mathbb{R}^3$ par :

$$J(\alpha_1, \alpha_2, x, y, z) = \frac{1}{2}x^2 + y^2 + z^2 + \frac{1}{6}\alpha_1.\alpha_2.x^6.y^6.z^6 + \alpha_1^3.\alpha_2^3.x.y.z^2.$$

Les notations sont celles de la proposition 3.1 . On prend $V = B(0,1)$ et $\Omega = B(0,1)$. Soit $\alpha_0 = (0,0)$, le point $x_0 = (0,0,0)$ est une solution du problème $\mathcal{P}(\alpha_0)$ qui vérifie la condition suffisante d'optimalité d'ordre 2 . C.à.d :

$$\forall x \in \mathbb{R}^3 , \; x \neq 0 \; , \; D^2_{xx} J(\alpha_0, x_0)(x,x) > 0 \; .$$

Donc :

$$\forall i \in \{1,2,3\} \; , \; g_i(\alpha_0, x_0) > 0 \; ,$$

et par conséquent , l'application h se réduit à f et $Q_{(\alpha_0,x_0)}$ se réduit au noyau de $Df(\alpha_0, x_0)$. Soit :

$$\cdot \; f_1(\alpha_1,\alpha_2,x,y,z) = \frac{\partial J}{\partial x}(\alpha_1,\alpha_2,x,y,z) = x + \alpha_1.\alpha_2.x^5.y^6.z^6 + \alpha_1^3.\alpha_2^3.y.z^2$$

$$\cdot \; f_2(\alpha_1,\alpha_2,x,y,z) = \frac{\partial J}{\partial y}(\alpha_1,\alpha_2,x,y,z) = 2y + \alpha_1.\alpha_2 x^6 y^5 z^6 + \alpha_1^3.\alpha_2^3.x.z^2$$

$$\cdot \; f_3(\alpha_1,\alpha_2,x,y,z) = \frac{\partial J}{\partial z}(\alpha_1,\alpha_2,x,y,z) = 2z + \alpha_1.\alpha_2 x^6 y^6 z^5 + 2\alpha_1^3.\alpha_2^3.x.y.z$$

$$\cdot \; Df(\alpha_0, x_0) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

et

$$\cdot \; Q_{(\alpha_0,x_0)} = \{ (\bar{\alpha}_1, \bar{\alpha}_2, \bar{x}, \bar{y}, \bar{z}) \in \mathbb{R}^5 / \; \bar{x} = 0 \, , \bar{y} = 0 \, , \bar{z} = 0 \} \; .$$

Détermination de A : Nous allons déterminer une constante $A>0$ telle que :

$$\forall (\alpha,x) \in V\times\Omega \ , \ \|D^2 f(\alpha,x)\| \leq A \ ,$$

une condition de croissance en découlera.

Rappelons que si $B=(b_{ij})$ est une matrice alors $\|B\|_2 = \left[\sup_{\lambda\in Sp({}^tB.B)} \lambda\right]^{1/2} \leq \left(\sum_{ij} b_{ij}^2\right)^{1/2}$. Donc pour tout (α,x) de

$V\times\Omega$, si on applique ce résultat aux matrices associées à $D^2 f_1(\alpha,x)$, $D^2 f_2(\alpha,x)$ et $D^2 f_3(\alpha,x)$, on obtient les majorations suivantes :

- $\|D^2 f_1(\alpha,x)\| \leq 103$
- $\|D^2 f_2(\alpha,x)\| \leq \frac{201}{2}$
- $\|D^2 f_3(\alpha,x)\| \leq 109$.

Donc : $\|D^2 f(\alpha,x)\| = \left(\sum_{i=1}^{3} \|D^2 f_i(\alpha,x)\|\right)^{1/2} \leq 180$,
et par conséquent :

$$\forall (\alpha,x) \in V\times\Omega \ , \ \|Df(\alpha,x) - Df(\alpha_0,x_0)\| \leq 180 \,\|(\alpha-\alpha_0, x-x_0)\| .$$

Détermination de M' : M' est réduit à $\mu(Df(\alpha_0,x_0))$.
Les valeurs propres de $Df(\alpha_0,x_0).{}^tDf(\alpha_0,x_0)$ sont 1 et 4, donc d'après la proposition 2.4.1.1 (cf ch II), $\mu(Df(\alpha_0,x_0))=1$.

Soit $M'=1$.

Soient α un paramètre voisin de α_0 et $X=(x,y,z)$ un point critique pour $\mathcal{P}(\alpha)$. D'après la proposition 3.1 ona :

$$d(\,(\alpha,x),\, Q_{(\alpha_0,x_0)}\,) \leqslant \frac{M'.A}{2} \;\|(\alpha,x)\|^2$$

soit :

$$\|x\| \leqslant 90\;(\,\|\alpha\|^2+\|x\|^2\,)\;,$$

donc pour tout k compris entre 0 et 1 , $\|x\| \leqslant \frac{90}{k}\;\|\alpha\|^2$ dès que $\|x\| \leqslant \frac{1-k}{90}$ ■

<u>Exemple</u> 2 : Soit J l'application numérique définie sur $\mathbb{R}\times\mathbb{R}^2$ par :

$$J(\alpha,x,y) = x^2+x^2.y+y^3+\alpha.x^4.y^4+\alpha.x^3.y^3+\alpha.x^2.y+\alpha^2xy^2+\alpha y.$$

Les notations sont celles de la proposition 3.1 . On prend $V=B(0,1)$ et $\Omega=B(0,1)$. Soit $\alpha_0=0$. Le point $x_0=(0,0)$ est un point critique pour $\mathcal{P}(\alpha_0)$ avec :

$$g_1(\alpha_0,x_0)>0 \quad \text{et} \quad g_2(\alpha_0,x_0)=0\;.$$

Donc : $h = (f, g_2)$ et $Q_a = \{ X \in \mathbb{R}\times\mathbb{R}^3 / Df(a)\cdot X = 0 ,\ Dg_2(a)\cdot X \geq 0 \}$.

Soit :

- $f_1(\alpha,x,y) = \frac{\partial J}{\partial x}(\alpha,x,y) = 2x + 2xy + 4\alpha x^3 y^4 + 3\alpha^3 x^2 y^3 + 2\alpha^2 xy + \alpha^2 y^2$

- $f_2(\alpha,x,y) = \frac{\partial J}{\partial y}(\alpha,x,y) = x^2 + 3y^2 + 4\alpha x^4 y^3 + 3\alpha^3 x^3 y^2 + \alpha^2 x^2 + 2\alpha^2 x\cdot y + \alpha$

- $g_2(\alpha,x,y) = \frac{\partial^2 J}{\partial x^2}(\alpha,x,y) \cdot \frac{\partial^2 J}{\partial y^2}(\alpha,x,y) - \left(\frac{\partial^2 J}{\partial x \partial y}(\alpha,x,y)\right)^2$

$$= (2 + 2y + 12\alpha x^2 y^4 + 6\alpha^3 x y^3 + 2\alpha^2 y) \times$$
$$(6y + 12\alpha x^4 y^2 + 6\alpha^3 x^3 y + 2\alpha^2 x)$$
$$- (2x + 16\alpha x^3 \cdot y^3 + 9\alpha^3 x^2 y^2 + 2\alpha^2 x + 2\alpha^2 y)^2$$

- $Dh(a) = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 12 \end{bmatrix}$

et

- $Q_a = \{ (\bar\alpha, \bar x, \bar y) \in \mathbb{R}^3 / \bar\alpha = 0 ,\ \bar x = 0 ,\ \bar y \geq 0 \}$

<u>Détermination de A</u> : Soit (α, x) un point de $V \times \Omega$, avec la même méthode de l'exemple 1, on obtient :

- $\| D^2 f_1(\alpha,x) \| \leq 153$
- $\| D^2 f_2(\alpha,x) \| \leq 162$
- $\| D^2 g_2(\alpha,x) \| \leq 204$.

Donc : $\|D^2h(\alpha,x)\| \leq 302$, et par conséquent :

$$\forall (\alpha,x) \in V\times\Omega \;,\; \|Dh(\alpha,x) - Dh(a)\| \leq 302\ \|(\alpha-\alpha_0, x-x_0)\| .$$

<u>Détermination de M'</u> : on a $|I_0|=1$, donc ,

$$M' = \mu\,(Dh(a))\cdot \|Dh(a)\| \cdot \sup_{[2]\subset I\subset[3]} \mu(Dh(a)_I) .$$

Les parties I qui vérifient $[2]\subset I\subset[3]$ sont :
$I_1 = \{1,2\}$ et $I_2 = \{1,2,3\}$ et les matrices associées sont :

$$Dh(a)_{I_1} = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ et } Dh(a)_{I_2} = Dh(a) = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 12 \end{bmatrix} .$$

. Les valeurs propres de $Dh(a)_{I_1} \cdot {}^tDh(a)_I$ sont 1 et 4
. Les valeurs propres de $Dh(a)\cdot{}^tDh(a)$ sont 1, 4 et 144 .

Donc d'après la proposition 2.4.1.1 (cf ch II) :

. $\mu\,(Dh(a)_I) = 1$
. $\mu\,(Dh(a)) = 1$
. $\mu\ \|Dh(a)\| = 12$.

D'où : $M' = 12$

Soient α un paramètre de V et $X = (x,y)$ un point critique pour $\mathcal{P}(\alpha)$. D'après la proposition 3.1 on a :

$$d((\alpha,x),Q_a) \leqslant \frac{12.302}{2} \cdot \|(\alpha,x)\|^2 .$$

Soit :

$$\begin{cases} 1 \leqslant 1812 \cdot (\alpha^2+x^2+y^2)^{1/2} & \text{si } (\alpha,x) \neq (\alpha_0,x_0) \text{ et } y \leq 0 \\ \sqrt{\alpha^2+x^2} \leqslant 1812 \cdot (\alpha^2+x^2+y^2) & \text{si } y > 0 \end{cases} .$$

Donc pour tout $k > 0$

$$\text{si } \begin{cases} 0 \neq \|x\| \leqslant \dfrac{1}{\sqrt{k^2+(1812)^2}} \\ \text{et} \\ y \leqslant 0 \end{cases} \qquad \text{alors } \|x\| \leqslant \frac{1812}{k} \|\alpha\| .$$

$$\text{si } \begin{cases} \|x\| \leqslant \dfrac{1}{\sqrt{k+1812}} \\ \text{et} \\ y > 0 \end{cases} \qquad \text{alors } \|x\| \leqslant \sqrt{\frac{1+1812\,\alpha^2-|\alpha|}{k}} \quad \blacksquare$$

Références bibliographiques :

[1] : p. Antoine et H. Zouaki, étude locale de l'ensemble des points critiques d'un problème d'optimisation paramètré, C.R. Acad. Sc. Paris, t. 310, Serie I, p. 587. 590. 1990.

[2] : J. p Penot et p. Terpolilli, cônes tangents et singularités, C.R. Acad. Sc. Paris, t. 296. 1983.

[3] : R. J. Magnus, on the local structure of the zero-set of a Banach space valued mapping, J. of func. anal, 22, 1976, p. 58-72

[4] : A. MAKKI NACIRI, Étude de l'approximation par son cône tangent d'une partie définie par des égalités et des inégalités. C.R. Acad. Sc. Paris (à paraître)

[5] : J. p Penot, on regularity conditions in mathematical programming, Mathematical programing studies, 19, 1982, p. 167. 199.

[6] : p. G. CIARRLET, introduction à l'analyse numérique matricielle et à l'optimisation (Masson).