



**HAL**  
open science

## Distance à un convexe d'un espace normé et caractérisation de points proximaux

Jean Jacques Moreau

► **To cite this version:**

Jean Jacques Moreau. Distance à un convexe d'un espace normé et caractérisation de points proximaux. 1969. hal-02162149

**HAL Id: hal-02162149**

**<https://hal.science/hal-02162149>**

Submitted on 21 Jun 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DISTANCE A UN CONVEXE D'UN ESPACE NORME  
ET CARACTERISATION DES POINTS PROXIMAUX

J.J. MOREAU

---

### 1. INTRODUCTION

L'évaluation de la distance d'un point  $z$  à une partie convexe (en particulier, un sous-espace vectoriel) d'un espace vectoriel normé, c'est à dire  $\inf\{\|x - z\| : x \in C\}$ , puis la caractérisation des éventuels points de  $C$  où cet inf est atteint (dits "points proximaux" de  $z$  dans  $C$  ou "projections" de  $z$  sur  $X$ ) constitue la formalisation commune d'un nombre considérable de problèmes dans des domaines très divers : théorie de l'approximation, optimisation, résolution de problèmes de physique mathématique par l'intermédiaire de principes extrémaux, etc... Aussi cette question a-t-elle donné lieu à de nombreux travaux ; le présent exposé ne prétend pas fournir une synthèse exhaustive de tous les résultats acquis dans cette voie, et la bibliographie qui l'accompagne ne saurait être complète.

Nous avons seulement voulu présenter les éléments de base sous une forme aussi maniable et explicative que possible. Le plus efficace nous a paru de fonder l'exposé sur le théorème "mini-max", ou plutôt "inf-max", que voici :

Soit  $C$  une partie convexe d'un espace vectoriel ; soit  $K$  une partie convexe compacte d'un espace vectoriel topologique ; soit  $(x, y) \mapsto f(x, y)$  une application de  $C \times K$  dans  $\mathbb{R}$ , convexe

par rapport à  $x$ , concave et s.c.s. par rapport à  $y$ . Alors

$$\inf_{x \in C} \max_{y \in K} f(x, y) = \max_{y \in K} \inf_{x \in C} f(x, y)$$

Cet énoncé peut être considéré comme un cas particulier de celui donné par K. FAN [1] (lequel invoque une forme affaiblie de convexité) ou comme un cas particulier de celui donné par l'auteur (Cf. MOREAU [2]) ; en fait, il suffit pour l'essentiel de ce qui suit de se limiter au cas particulier où  $f$  est bilinéaire : la priorité de l'énoncé revient alors à H. KNESER [1].

On invoquera le fait que la boule unité du dual d'un espace normé est faiblement compacte.

De cette manière on atteint rapidement les résultats essentiels. Toutefois, celui qui voudrait retracer l'itinéraire logique suivi depuis les fondements de la théorie des espaces vectoriels topologiques pourra trouver le chemin assez détourné. De fait, si on se reporte à la démonstration du théorème "inf-max", ainsi qu'elle ressort de MOREAU [1], [2], on constatera que l'essentiel n'est pas proprement la compacité de  $K$ , mais son équicontinuité relativement à une certaine dualité ; justement ici, une telle équicontinuité est bien le caractère principal de la boule unité du dual de notre espace normé. C'est pourquoi nous terminons l'exposé en reprenant la question par le moyen de la théorie de la dualité et de la sous-différentiabilité des fonctionnelles convexes (Cf. pour un exposé systématique, MOREAU [3]). On fait par là appel à des techniques moins largement connues, mais nous espérons ainsi illustrer la grande efficacité heuristique de ces méthodes. Nous en profitons pour adopter dans ces deux derniers paragraphes un point de vue plus général : au lieu d'une norme sur l'espace  $X$  nous considérons, à l'exemple de P.J. LAURENT [1],

une fonction sous-linéaire positive quelconque (non nécessairement finie), soit  $a$ .

L'artifice technique de départ consiste à remarquer que la recherche de l'inf de la fonction  $a$  sur l'ensemble  $C$  et la caractérisation des éventuels points de  $C$  où cet inf est atteint équivaut à la même étude, effectuée relativement à l'espace  $X$  entier, pour la fonction  $a + \psi_C$  (on note  $\psi_C$  la fonction indicatrice de  $C$ , qui prend par définition la valeur 0 sur  $C$  et  $+\infty$  en dehors ; c'est une fonction convexe si et seulement si l'ensemble  $C$  est convexe). L'idée d'appliquer la dualité à l'étude de l'inf d'une somme de fonctions convexes était déjà présente chez W. FENCHEL [2] ; il s'agissait pour cet auteur de fonctions convexes finies, définies sur des parties convexes de  $\mathbb{R}^n$  ; l'adaptation du résultat de FENCHEL aux e.v.t. de dimension quelconque a été menée à bien par R.T. ROCKAFELLAR [1] (pour obtenir une écriture légèrement plus harmonieuse, ces auteurs préfèrent d'ailleurs considérer la différence entre une fonction convexe et une fonction concave).

Par souci de clarté, on s'est limité tout au long de l'étude qui va suivre au cas où le point  $z$  est l'origine de  $X$  ; le cas général s'y ramène par translation et le soin de transcrire les énoncés est laissé au lecteur.

L'auteur remercie Monsieur H. ARIS, Assistant au C.S.U. de Perpignan, pour la très sérieuse enquête bibliographique qu'il a consacrée à la question.

## 2. EXPRESSION "inf-max" DE LA DISTANCE

Dans tout ce qui suit,  $X$  désigne un espace vectoriel réel, muni d'une norme notée  $\|\cdot\|$  ; on appelle  $Y$  son dual topologique,

classiquement muni d'une norme qu'on notera aussi  $\|\cdot\|$  : pour tout  $y \in Y$

$$\|y\| = \sup \{ \langle x, y \rangle \in \mathbb{R} ; x \in X, \|x\| \leq 1 \} .$$

Soit

$$B = \{ y \in Y : \|y\| \leq 1 \}$$

la boule unité fermée de  $Y$  ; il est bien connu que, pour tout  $x \in X$ , on a inversement

$$(2.1) \quad \|x\| = \sup_{y \in B} \langle x, y \rangle$$

et que ce "sup" est, en fait, un "max" .

Soit alors  $C$  une partie convexe non vide de  $X$  . La distance de l'origine à  $C$  est, compte tenu de (2.1),

$$(2.2) \quad d = \inf_{x \in C} \|x\| = \inf_{x \in C} \max_{y \in B} \langle x, y \rangle .$$

En invoquant le théorème "inf-max" et la compacité de  $B$  pour la topologie faible  $\sigma(Y, X)$  on en tire

$$(2.3) \quad d = \max_{y \in B} \inf_{x \in C} \langle x, y \rangle .$$

Variante.

En utilisant le fait que la symétrie dans  $\mathbb{R}$  transmue les "inf" en "sup", puis le fait que  $B$  est un ensemble symétrique dans  $Y$ , on transforme (2.3) en

$$(2.4) \quad -d = \min_{y \in B} \sup_{x \in C} \langle x, y \rangle = \min_{y \in B} \varphi_C(y) ,$$

où  $\varphi_C$  désigne la fonction d'appui de  $C$  .

3. CARACTERISATION DES POINTS PROXIMAUX PAR UNE PROPRIETE DE COL

Un point proximal de l'origine dans  $C$  est un point  $x_0 \in C$  tel que  $\|x_0\| = d$ .

PROPOSITION. Un point  $x_0 \in C$  et proximal de l'origine dans  $C$  si et seulement si on peut lui associer  $y_0 \in B$  tel que  $(x_0, y_0)$  soit un col (minimal en  $x$ , maximal en  $y$ ) de l'application  $(x, y) \mapsto \langle x, y \rangle$  de  $C \times B$  dans  $\mathbb{R}$ , c'est à dire si, pour tout  $x \in C$  et tout  $y$  dans  $B$ , on a

$$(3.1) \quad \langle x_0, y \rangle \leq \langle x_0, y_0 \rangle \leq \langle x, y_0 \rangle .$$

Suffisance : Supposons  $x_0 \in C$  et  $y_0 \in B$  assurant la propriété de col en question ; d'après la première inégalité (3.1) on a

$$(3.2) \quad \|x_0\| = \sup_{y \in B} \langle x_0, y \rangle = \langle x_0, y_0 \rangle .$$

Pour tout  $x \in C$ , la deuxième inégalité donne alors

$$\|x\| = \sup_{y \in B} \langle x, y \rangle \geq \langle x, y_0 \rangle \geq \langle x_0, y_0 \rangle = \|x_0\| ,$$

ce qui prouve que  $x_0$  est proximal.

Nécessité : Soit  $x_0 \in C$  tel que  $\|x_0\| = d$  ; d'après (2.1) et (2.3) on a

$$\max_{y \in B} \langle x_0, y \rangle = \max_{y \in B} \inf_{x \in C} \langle x, y \rangle .$$

Soit  $y_0 \in B$  un point où le max du second membre est atteint, de sorte que

$$\max_{y \in B} \langle x_0, y \rangle = \inf_{x \in C} \langle x, y_0 \rangle ,$$

ce qui implique :

$$\forall x \in C, \forall y \in B : \langle x_0, y \rangle \leq \langle x, y_0 \rangle .$$

En faisant successivement  $x = x_0$  et  $y = y_0$ , on en tire les deux inégalités (3.1).

REMARQUE 1. Si les conditions précédentes sont satisfaites, (3.2) montre que

$$\|x_0\| = \langle x_0, y_0 \rangle = d .$$

Mais, essentiellement,

$$\langle x_0, y_0 \rangle \leq \|x_0\| \|y_0\| ,$$

donc

$$\|x_0\| \leq \|x_0\| \|y_0\| .$$

Si  $d \neq 0$ , c'est à dire si l'origine n'est pas un point adhérent à  $C$  on en tire (vu que  $\|y_0\| \leq 1$ )

$$\|y_0\| = 1 .$$

Par contre, si  $d = 0$ , de sorte qu'un point proximal  $x_0$  ne peut être que l'origine elle-même, un  $y_0 \in B$  associé à ce point proximal n'a pas nécessairement une norme égale à 1. En ce cas, en effet, l'inégalité (3.1) impliquera

$$\forall x \in C : \langle x, y_0 \rangle \leq 0 .$$

Si  $y_0 \neq 0$  cela signifie que l'hyperplan fermé

$$\{x \in X : \langle x, y_0 \rangle = 0\}$$

est hyperplan d'appui pour  $C$  à l'origine. On sait qu'il n'existe pas nécessairement de tel hyperplan.

REMARQUE 2. Un  $y_0 \in B$  satisfait la condition (3.2) si et seulement si

$$\forall x \in C : \langle x, y_0 \rangle \geq d ,$$

ce qui définit un ensemble  $K$  convexe, faiblement compact puisque c'est l'intersection de  $B$  avec l'ensemble convexe (faiblement) fermé

$$\{y \in Y : \forall x \in C, \langle x, y \rangle \geq d\} .$$

Il est important de constater que s'il y a plusieurs points proximaux de l'origine dans  $C$  (noter qu'ils forment un ensemble convexe) il leur correspond ce même ensemble  $K$  de points  $y_0$  associés.

COROLLAIRE. (RUBINSTEIN [1]).  $x_0 \in C$  non nul est point proximal de l'origine si et seulement si il existe  $y_0 \in Y$  tel que

$$\|y_0\| = 1 \quad (\text{ou aussi bien } \|y_0\| \leq 1)$$

$$\langle x_0, y_0 \rangle = \|x_0\|$$

$$\forall x \in C : \langle x - x_0, y_0 \rangle \geq 0$$

et on peut prendre le même  $y_0$  pour tous les éventuels points proximaux de l'origine dans  $C$ .

#### 4. HYPERPLANS SEPARATEURS ET ENSEMBLES CO-POLAIRES

Tout hyperplan fermé de  $X$ , ne passant pas par l'origine peut être, d'une seule manière, écrit sous la forme

$$\Pi = \{x \in X : \langle x, y \rangle = 1\} .$$

Nous disons que  $y \in Y$  est le pôle de l'hyperplan  $\Pi$ . On trouve aisément (voir Appendice) que la distance de l'origine à  $\Pi$  est égale à  $1/\|y\|$ ; cela généralise une situation bien connue de géométrie élémentaire : si  $X$  est un espace vectoriel euclidien, identifié à son dual,  $y$  est le pôle de  $\Pi$  par rapport à la sphère unité.

L'hyperplan  $\Pi$  sépare  $C$  de l'origine si et seulement si

$$(4.1) \quad \forall x \in C : \quad \langle x, y \rangle \geq 1 .$$

Nous poserons

$$\gamma(y) = \inf_{x \in C} \langle x, y \rangle = -\varphi_C(-y)$$

( $\varphi_C$  : fonction d'appui de  $C$ ) ; étant l'enveloppe inférieure d'une famille de fonctions linéaires continues, la fonction numérique  $\gamma$ , partout définie sur  $Y$ , est sur-linéaire (i.e. concave et positivement homogène) et s.c.s. (pour toutes les topologies compatibles avec la dualité).

Comme on suppose  $C$  non vide,  $\gamma(0) = 0$ .

La condition (4.1) se reformule comme suit :

L'hyperplan  $\Pi$  sépare  $C$  de l'origine si et seulement si son pôle  $y$  appartient à l'ensemble (convexe, fermé, ne contenant pas l'origine de  $Y$ )

$$(4.2) \quad D = \{y \in Y : \gamma(y) \geq 1\} ,$$

que nous appellerons ensemble co-polaire de  $C$ .

La fonction  $\gamma$  prend quelque part des valeurs  $> 0$  si et seulement si  $D$  est non vide, c'est à dire si et seulement si il

existe des hyperplans fermés séparant strictement  $C$  de l'origine, c'est à dire si et seulement si l'origine n'est pas adhérente à  $C$ , autrement dit si et seulement si  $d \neq 0$  (en ne supposant pas  $C$  convexe, on a  $D \neq \emptyset$  si et seulement si l'enveloppe convexe fermée de  $C$  ne contient pas l'origine).

Bref

$$(4.3) \quad D = \{y \in Y : \forall x \in C, \langle x, y \rangle \geq 1\} .$$

Cet ensemble est donc une intersection de demi espaces fermés ne contenant pas l'origine de  $Y$ .

En faisant jouer de manière symétrique la dualité entre  $X$  et  $Y$  on peut définir l'ensemble co-polaire de  $D$

$$C' = \{x \in X : \forall y \in D, \langle x, y \rangle \geq 1\} .$$

$C'$  est l'intersection des demi espaces fermés ne contenant pas l'origine et contenant  $C$  (ou "enveloppe co-polaire" de  $C$ ). Les ensembles  $C'$  et  $D$  sont mutuellement co-polaires (pour la généralité, on notera que si l'un de ces deux ensembles est vide, l'autre est la partie pleine de l'espace vectoriel correspondant ; si, par exemple, l'origine est dans l'enveloppe convexe fermée de  $C$ , la famille des demi espaces fermés contenant  $C$  et ne contenant pas l'origine est vide : l'enveloppe co-polaire de  $C$  étant intersection d'une famille vide de parties de  $X$  est bien  $X$  entier).

L'usage de la fonction  $\gamma$  donne à (2.3) la forme

$$(4.4) \quad d = \max_{y \in B} \gamma(y)$$

Pour invoquer le résultat de l'Appendice on introduira la fonction

$$\gamma^+(y) = \begin{cases} \gamma(y) & \text{si } \gamma(y) \geq 0 \\ 0 & \text{si } \gamma(y) \leq 0 \end{cases} .$$

Si  $D \neq \emptyset$  (c'est à dire  $d \neq 0$ ) (4.4) équivaudra à

$$\begin{aligned} d &= \max_{y \in B} \gamma^+(y) \\ &= \sup \{ \gamma^+(y) : \|y\| \leq 1 \} . \end{aligned}$$

D'après l'Appendice cette quantité est l'inverse de

$$\inf \{ \|y\| : \gamma^+(y) \geq 1 \} = \inf_{y \in D} \|y\| ,$$

c'est à dire la distance de l'origine de  $Y$  à  $D$ . Dans le cas  $D \neq \emptyset$ , la distance en question est  $+\infty$  et  $d = 0$ ; de même si  $C$  est vide,  $d = +\infty$  tandis que  $D = Y$ ; on peut donc énoncer

PROPOSITION. La distance de l'origine à une partie convexe quelconque de  $X$  et la distance de l'origine à son copolaire dans  $Y$  sont inverses l'une de l'autre.

Il en résulte, en particulier, que  $d$  est aussi bien la distance de l'origine à l'"enveloppe copolaire"  $C'$  de  $C$ .

REMARQUE.

La proposition précédente s'écrit encore

$$(4.5) \quad d = \sup_{y \in D} \frac{1}{\|y\|} ,$$

ce qui donne lieu à l'interprétation très naturelle que voici : Les  $y \in D$  définissent les demi espaces fermés contenant  $C$  et ne contenant pas  $0$  :  $1/\|y\|$  est, comme on l'a dit, la distance de  $0$  au demi-espace correspondant (où à la frontière de ce demi-espace) ; il est assez intuitif que le sup de ces distances est la distance de  $0$  à  $C$  .

5. USAGE DES POINTS EXTREMAUX DE LA BOULE UNITE DE  $Y$  .

Notons d'abord que, la forme linéaire  $y \mapsto \langle x, y \rangle$  étant faiblement continue sur  $Y$ , son max sur  $B$ , ensemble convexe faiblement compact est égal aux max de cette forme sur l'ensemble  $\ddot{B}$  des points extrémaux (ou "profil" de  $B$ ) . En effet si  $m$  est ce max, l'hyperplan fermé  $\{y \in Y : \langle x, y \rangle = m\}$  (hyperplan propre si  $x \neq 0$  ; si  $x = 0$  le résultat trivial) est d'appui pour  $B$  ; son intersection avec  $B$  est un convexe faiblement compact lequel possède nécessairement au moins un point extrémal ; il est immédiat que ce point est aussi extrémal pour  $B$  .

Donc (2.1) se raffine en

$$(3.1) \quad \|x\| = \max_{y \in \ddot{B}} \langle x, y \rangle$$

d'où

$$d = \inf_{x \in C} \max_{y \in \ddot{B}} \langle x, y \rangle$$

Mais ici, on ne peut plus espérer intervertir inf et max .

On peut, d'autre part, invoquer le théorème de Krein-Milman moyennant lequel, dans un énoncé, la clause:  $y_0 \in B$  pourra être remplacée par :  $y_0$  appartient à l'enveloppe convexe fermée de  $\ddot{B}$  . Cela

sera surtout intéressant si l'on peut se ramener à un ensemble de dimension finie ; en effet moyennant le théorème de Caratheodory, le théorème de Krein-Milman, pour les compacts de dimension finie, se raffine en :  $y$  appartient à  $K$ , convexe compact de dimension finie  $p$ , si et seulement si il existe  $p+1$  points  $y_i \in \overset{\circ}{\overset{\circ}{K}}$  et  $p+1$  réels  $\alpha_i \in [0, 1]$  avec  $\sum \alpha_i = 1$  tels que

$$y = \sum \alpha_i y_i .$$

On va appliquer cette idée au cas où  $C$  étant, dans  $X$ , un ensemble de dimension finie, il est contenu dans un sous-espace vectoriel  $E_n$  (si  $E_n$  est le plus petit sous-espace vectoriel contenant  $C$ , la dimension affine de  $C$  est  $n$  ou  $n-1$ ). L'espace  $E_n$  étant muni de la norme restriction de celle de  $X$ , appelons  $F_n$  son dual. Appliquons le corollaire du § 3 aux espaces  $E_n$  et  $F_n$  en remarquant que si  $f_0 \in F_n$  vérifie la condition

$$\langle x_0, f_0 \rangle = \|x_0\| ,$$

ce point appartient à l'ensemble  $\{f : \langle x_0, f \rangle = \|x_0\|\}$  lequel est un hyperplan d'appui  $\Pi$  de la boule unité  $B_n$  de  $F_n$ . L'intersection de  $\Pi$  avec  $B_n$  est un compact de dimension affine  $n-1$  au plus, auquel on applique le théorème de Krein-Milman en remarquant d'autre part que tout point extrémal de cette intersection est aussi extrémal pour  $B_n$ . Il vient ainsi :  $x_0 \in C$ , non nul, est point proximal de l'origine (dans  $E_n$ , donc aussi bien dans  $X$ ) si et seulement si il existe  $n$  points extrémaux  $f_i$  de la boule unité  $B_n$  de  $E_n$  et  $n$  réels  $\alpha_i \in [0, 1]$ , de somme 1, tels que

$$\sum_i \alpha_i \langle x_0, f_i \rangle = \|x_0\|$$

$$\forall x \in C : \sum_i \alpha_i \langle x - x_0, f_i \rangle \geq 0$$

Remarquons maintenant que  $F_n$  peut être identifié à l'espace quotient de  $Y$  par l'orthogonal de  $E_n$  ; cela définit une application linéaire  $P$ , "projection" de  $Y$  sur  $F_n$ . Cette application linéaire est continue de  $Y$  (muni de la topologie faible  $\sigma(Y, X)$ ) dans  $F_n$  et  $B_n = P(B)$ . Il est élémentaire que tout point extrémal de l'image d'un convexe compact par une application linéaire continue est l'image d'au moins un point extrémal de ce compact. Bref :

PROPOSITION (RUBINSTEIN [1]). Si  $C$  est contenu dans un sous-espace de dimension  $n$  de  $X$ , un point  $x_0 \in C$ , non nul, est proximal de l'origine si et seulement si il existe  $n$  points extrémaux de la boule unité  $B$  de  $Y$ , soient  $y_i$ , et  $n$  réels  $\alpha_i \in [0, 1]$  de somme 1 tels que

$$\sum_i \alpha_i \langle x_0, y_i \rangle = \|x_0\|$$

$$\forall x \in C : \sum_i \alpha_i \langle x - x_0, y_i \rangle \geq 0$$

(noter en effet que si  $f_i = P(y_i)$  et  $x - x_0 \in E_n$  on a  $\langle x - x_0, f_i \rangle = \langle x - x_0, y_i \rangle$ ).

Pour l'application d'idées de cet ordre à la caractérisation des points proximaux d'un point donné d'un espace normé dans un sous-espace vectoriel de dimension finie, voir LAURENT [1] : cela fournit notamment des généralisations du théorème classique de Tchebycheff.

## 6. USAGE DE LA THEORIE DES FONCTIONS DUALES

Dans ce qui suit,  $X$  et  $Y$  sont deux espaces vectoriels mis en dualité par une forme bilinéaire notée  $\langle \cdot, \cdot \rangle$ . Soient  $A \subset X$  et  $B \subset Y$  deux ensembles mutuellement polaires (donc convexes contenant l'origine, fermés pour les topologies compatibles avec la dua-

lité). Soient  $a$  et  $b$  leurs jauges respectives ; on peut dire aussi bien que  $a$  est la fonction d'appui de  $B$  :

$$a(x) = \sup_{y \in B} \langle x, y \rangle$$

et  $b$  la fonction d'appui de  $A$ . L'exemple initial où  $X$  était un espace normé et  $Y$  son dual, consiste à prendre pour  $a$  et  $b$  les normes respectives de ces deux espaces ;  $A$  et  $B$  sont alors les boules unités respectives.

$a$  et  $b$  sont des fonctions sous-linéaires  $\geq 0$ , semi-continues inférieurement pour les topologies compatibles avec la dualité. Il y a au départ symétrie complète entre les rôles des deux espaces : formulons quelques propriétés pour  $a$  (Cf. MOREAU [3]).

$a$  est une fonction paire si et seulement si  $B$  est une partie symétrique de  $X$  ;  $a$  est partout finie si et seulement si  $A$  est absorbant, ce qui équivaut à  $B$  borné dans  $Y$  (pour les topologies compatibles avec la dualité) ;  $a$  est strictement positive hors de l'origine si et seulement si  $A$  ne contient pas de demi-droite, ce qui revient à dire que  $B$  n'admet pas d'hyperplan d'appui (fermé pour les topologies compatibles avec la dualité) à l'origine (i.e. il n'est pas contenu dans l'un des demi-espaces que définirait un tel hyperplan). Bref  $a$  est une norme si et seulement si toutes les conditions précédentes sont remplies.

$a$  est continue à l'origine (donc partout dans  $X$ ) pour une topologie compatible avec la dualité (donc a fortiori pour la topologie de Mackey  $\tau(X, Y)$ ) si et seulement si  $B$  est compact dans  $Y$  pour la topologie faible  $\sigma(Y, X)$  ;  $a$  est continue pour la topologie forte de  $X$  (topologie de la convergence uniforme dans les

parties de  $Y$  qui sont bornées par rapport à la dualité) si et seulement si elle est partout finie.

Soit  $C$  une partie convexe de  $X$  ; par analogie avec les paragraphes précédents, nous cherchons des évaluations de

$$(6.1) \quad d = \inf_{x \in C} a(x) .$$

En introduisant la fonction indicatrice  $\psi_C$  (qui vaut zéro sur  $C$  et  $+\infty$  ailleurs) cela s'écrit aussi

$$(6.2) \quad d = \inf_{x \in X} [\psi_C(x) + a(x)] .$$

Une des voies possibles pour faire entrer cette expression dans des considérations de dualité est d'invoquer l'opération d'inf-convolution de deux fonctions numériques  $f$  et  $g$  définies sur  $X$  :

$$(6.3) \quad (f \nabla g)(x) = \inf_{u \in X} [f(u) + g(x - u)]$$

En introduisant la fonction indicatrice de l'ensemble  $-C$ , (6.2) prend aussi la forme

$$d = (\psi_{-C} \nabla a)(0) .$$

La fonction  $\psi_{-C} \nabla a$  est convexe si la fonction  $a$  et l'ensemble  $C$  le sont ; mais elle n'appartient pas nécessairement à l'ensemble, noté  $\Gamma(X, Y)$ , des fonctions auxquelles s'applique directement la théorie de la dualité (fonctions convexes, s.c.i., à valeurs dans  $]-\infty, +\infty]$ ). Toutefois, si  $\psi_{-C}$  et  $a$  vérifient des conditions telles que la fonction  $\psi_{-C} \nabla a$  soit s.c.i. à l'origine, alors sa valeur en ce point sera aussi bien celle de sa régularisée s.c.i., soit  $(\psi_{-C} \nabla a)^{**}$ . Or on montre que cette régularisée a pour

fonction duale la somme des fonctions duales  $\psi_{-C}^*$  et  $a^*$  de  $\psi_C$  et  $a$  ; d'après la définition même des fonctions polaires, cela donne

$$\begin{aligned} (\psi_{-C} \nabla a)^{**}(0) &= - \inf_{y \in Y} [\psi_{-C}^*(y) + a^*(y)] \\ &= - \inf_{y \in B} \psi_{-C}^*(y) \end{aligned}$$

car  $a^*$  n'est autre que la fonction indicatrice  $\psi_B$  de l'ensemble  $B$  (rappelons que  $\psi_{-C}^*$  est la fonction d'appui de l'ensemble  $-C$ ) .

Si  $f$  et  $g$  sont des fonctions convexes une condition suffisante usuelle pour que la fonction  $f \nabla g$  soit continue au point  $0$ , donc a fortiori s.c.i. en ce point, est la suivante : existence d'un point  $v_0 \in X$  en lequel  $g$  est finie et continue avec, d'autre part  $f(-v_0)$  fini. La définition (6.3) montre en effet que la fonction (convexe)  $f \nabla g$  est majorée par la fonction

$$x \longmapsto f(-v_0) + g(x + v_0)$$

(prendre  $u = -v_0$ ) laquelle est finie et continue à l'origine. Si l'on songe d'autre part que la continuité de la fonction indicatrice  $\psi_{-C}$  en un point signifie simplement que ce point est intérieur à l'ensemble  $-C$ , il vient :

PROPOSITION. Si, pour une topologie compatible avec la dualité, il existe un point de  $C$  en lequel  $a$  est finie et continue (resp. un point intérieur de  $C$  en lequel  $a$  est finie), on a

$$(6.4) \quad d = - \inf_{y \in B} \psi_{-C}^*(y)$$

et l'inf ici écrit est un minimum.

La dernière assertion résulte de ce que la fonction convexe  $\psi_{-c} \nabla a$ , étant finie et continue à l'origine, possède une minorante affine continue "exacte" en ce point ; selon la théorie des fonctions duales, cela signifie l'existence d'un  $y_0 \in Y$  qui est conjugué de l'origine et cela s'écrit

$$(\psi_{-c} \nabla a)(0) + (\psi_{-c}^* + a^*)(y_0) = \langle 0, y_0 \rangle$$

c'est à dire (puisque  $a^* = \psi_B$ ) qu'il existe  $y_0 \in B$  tel que

$$(\psi_{-c} \nabla a)(0) = - \psi_{-c}^*(y_0) .$$

Un autre type de condition assurant que la fonction  $\psi_{-c} \nabla a$  appartient à  $\Gamma(X, Y)$ , de sorte qu'elle est égale à  $(\psi_{-c} \nabla a)^{**}$ , est celui où les deux fonctions  $\psi_{-c}$  et  $a$  sont s.c.i. et l'une d'entre elles inf-compacte ; on obtient ainsi :

PROPOSITION. Si, pour une topologie compatible avec la dualité (donc, a fortiori, pour la topologie faible  $\sigma(X, Y)$ ) l'ensemble convexe  $C$  est compact, ou encore si  $C$  est fermé et  $A$  compact, on a l'égalité (6.4). On ne peut affirmer en ce cas que l'inf de cette égalité soit un minimum ; par contre il en est ainsi pour l'inf dans la définition (6.1) de  $d$ .

Cette dernière assertion résulte du fait que, sous les hypothèses faites, l'inf-convolution  $\psi_{-c} \nabla a$  est exacte (cf. MOREAU [3], § 4.e).

En fait les hypothèses de cette proposition équivalent, d'après la théorie de la dualité (Cf. MOREAU [3], § 8.f) à ce que l'une des fonctions  $\psi_{-c}^*$  ou  $a^* = \psi_B$  est continue à l'origine pour une topologie compatible avec la dualité (donc a fortiori pour la topologie

de Mackey  $\tau(Y, X)$  . On pourrait obtenir un résultat plus général en remplaçant l'origine de  $Y$  par un  $y_0 \in Y$  , "pente" par rapport à laquelle l'une des deux fonctions  $\psi_c$  et  $a$  serait inf-compacte. Dans ces conditions les deux propositions du présent paragraphe apparaissent comme duales l'une de l'autre.

Résumons en constatant que les diverses hypothèses envisagées dans les deux propositions ci-dessus assurent toutes la relation

$$\inf_{x \in C} \sup_{y \in B} \langle x, y \rangle = \sup_{y \in B} \inf_{x \in C} \langle x, y \rangle .$$

Comme au § 4 on pourra introduire l'ensemble  $D$  copolaire de  $C$  :

$$\begin{aligned} D &= \{y \in Y : \forall x \in C, \langle x, y \rangle \geq 1\} \\ &= \{y \in Y : -\psi_{-c}^*(y) \geq 1\} . \end{aligned}$$

Vu que

$$B = \{y \in Y : b(y) \leq 1\} ,$$

le même mode de raisonnement qu'alors montre que, moyennant l'une des hypothèses formulées dans les deux propositions précédentes, l'inf de  $a$  sur  $C$  et l'inf de  $b$  sur  $D$  sont des nombres inverses.

## 7. SOUS-DIFFERENTIELS ET CARACTERISATION DES POINTS PROXIMAUX

Par analogie avec les paragraphes antérieurs, nous qualifions de proximal un point  $x_0 \in C$  tel que

$$(7.1) \quad a(x_0) = d .$$

Il s'agit donc d'un point de  $X$  où la fonction  $\psi_c + a$  atteint son

inf. Cela peut s'exprimer en disant que le sous-différentiel de la fonction  $\psi_C + a$  en ce point contient le zéro de  $Y$  :

$$(7.2) \quad 0 \in \partial(\psi_C + a)(x_0) .$$

Supposons que les fonctions  $\psi_C$  et  $a$  satisfassent des conditions assurant pour tout  $x \in X$  l'égalité d'ensembles :

$$(7.3) \quad \partial(\psi_C + a)(x) = \partial \psi_C(x) + \partial a(x) ;$$

alors la condition (7.1) équivaut à l'existence d'un  $y_0 \in Y$  tel que

$$(7.4) \quad y_0 \in \partial a(x_0)$$

$$(7.5) \quad -y_0 \in \partial \psi_C(x_0) .$$

Or, comme  $a$  et  $\psi_B$  sont des fonctions duales, (7.4) s'écrit aussi bien

$$y_0 \in B , \quad a(x_0) = \langle x_0 , y_0 \rangle ,$$

tandis que (7.5) s'écrit

$$x_0 \in C \text{ et } (\forall x \in C : \langle x - x_0 , y_0 \rangle \geq 0) .$$

Une condition suffisante pour que l'on ait l'égalité (7.2) est, comme au paragraphe 6, l'existence d'un point de  $X$  où l'une des deux fonctions est finie, l'autre finie et continue (Cf. ROCKAFELLAR [1], MOREAU [3], § 10.d) ; moyennant cela on obtient un résultat analogue au corollaire du § 3 :

PROPOSITION. S'il existe un point de  $C$  où la fonction  $a$  est finie et continue pour une topologie compatible avec la dualité (resp. s'il existe un point intérieur de  $C$  en lequel la fonction  $a$  est finie) ,

un point  $x_0 \in C$  assure l'égalité (7.1) si et seulement si il existe  
 $y_0 \in B$  tel que

$$a(x_0) = \langle x_0, y_0 \rangle$$

$$\forall x \in C : \langle x - x_0, y_0 \rangle \geq 0 .$$

A cela on peut adjoindre des remarques semblables à celles  
 du § 3:

1°) On a essentiellement (Cf. MOREAU [3] , § 14.b) :

$$(7.6) \quad \forall x_0 \in X, \forall y_0 \in Y : a(x_0) b(y_0) \geq \langle x_0, y_0 \rangle .$$

Donc, à condition que  $d \neq 0$ , les conditions précédentes impliqueront  
 $b(y_0) \geq 1$  c'est à dire  $b(y_0) = 1$  (puisque  $b$  est la jauge de  $B$ ).

2°) L'ensemble des  $y_0 \in Y$  que les conditions précédentes asso-  
 cient à un éventuel point proximal  $x_0$  est défini par

$$(7.7) \quad y_0 \in B$$

$$(7.8) \quad \langle x_0, y_0 \rangle = d$$

$$(7.9) \quad \forall x \in C : \langle x, y_0 \rangle \geq d .$$

Mais, vu (7.6), la condition (7.9) implique

$$a(x_0) b(y_0) \geq d$$

donc, toujours dans l'hypothèse que  $d = a(x_0)$  n'est pas nul elle im-  
 plique  $b(y_0) = 1$ , d'où  $d \geq \langle x_0, y_0 \rangle$ , d'où (7.8). L'ensemble de  
 ces  $y_0$  est donc simplement défini par les conditions (7.7) et (7.9) les-  
 quelles n'invoquent pas  $x_0$ , mais seulement  $d$ . Si de la sorte il

il y a plusieurs points proximaux  $x_0$  (noter qu'ils forment un ensemble convexe) il correspond à chacun d'eux le même ensemble (convexe) de points  $y_0$ .

Enfin la proposition précédente fournit une propriété de col semblable à celle du § 3 :

PROPOSITION. Sous les hypothèses de la Proposition précédente, un point  $x_0 \in C$  est proximal si et seulement si on peut lui associer  $y_0 \in B$  tel que pour tout  $x$  dans  $C$  et tout  $y$  dans  $B$  on ait :

$$(7.10) \quad \langle x_0, y \rangle \leq \langle x_0, y_0 \rangle \leq \langle x, y_0 \rangle .$$

En effet, supposons (7.10) réalisée : l'inégalité de gauche donne

$$a(x_0) = \sup_{y \in B} \langle x_0, y \rangle = \langle x_0, y_0 \rangle$$

et celle de droite, pour tout  $x \in C$ ,

$$a(x) = \sup_{y \in B} \langle x, y \rangle \geq \langle x, y_0 \rangle \geq \langle x_0, y_0 \rangle ,$$

ce qui montre que  $x_0$  est proximal.

Réciproquement, supposons  $x_0 \in C$  proximal ; la Proposition de tout à l'heure lui associe  $y_0 \in B$  tel que, pour tout  $x \in C$ , on ait

$$a(x_0) = \langle x_0, y_0 \rangle \leq \langle x, y_0 \rangle$$

et, pour tout  $y$  dans  $B$ ,

$$\langle x_0, y \rangle \leq \sup_{y \in B} \langle x_0, y \rangle = a(x_0)$$

ce qui prouve (7.10).

Notons, pour terminer, que des conditions assurant la propriété (7.3) d'addition des sous-différentiels pour des fonctions convexes s.c.i. ont été données par C. LESCARRET [1], dans la ligne des travaux de G. CHOQUET [1] sur les ensembles convexes faiblement complets. On en tirerait d'autres hypothèses suffisantes pour la validité des critères de proximalité formulés dans ce paragraphe. De même, un cas d'addition des sous-différentiels a été donné par l'auteur (MOREAU [4]) : il s'appliquera ici si la fonction  $a$  est faiblement différentiable (au moins hors de l'origine), ce qui suppose des propriétés de "rotondité" de l'ensemble  $B$ .

## 8. APPENDICE

Soit  $X$  un ensemble dans lequel opère le groupe multiplicatif des réels strictement positifs, c'est à dire qu'est défini, pour tout  $\lambda \in ]0, +\infty[$  et tout  $x \in X$ , l'élément  $\lambda x \in X$ , avec les règles de calcul

$$1 x = x$$

$$\lambda(\mu x) = (\lambda \mu) x .$$

Soient  $f$  et  $g$  deux applications de  $X$  dans  $[0, +\infty]$  positivement homogènes de degré 1, c'est à dire telles que, pour tout  $x \in X$  et tout  $\lambda \in ]0, +\infty[$ ,

$$f(\lambda x) = \lambda f(x)$$

et même propriété pour  $g$ .

Alors les deux nombres

$$\alpha = \inf \{f(x) : g(x) \geq 1\}$$

et

$$\beta = \sup \{g(x) : f(x) \leq 1\}$$

sont inverses (avec la convention que  $0$  et  $+\infty$  sont inverses).

D'ailleurs  $\alpha$  est aussi égal à

$$\alpha' = \inf \{f(x) : g(x) > 1\}$$

et  $\beta$  à

$$\beta' = \sup \{g(x) : f(x) < 1\} .$$

L'inf et le sup sont pris dans  $[0, +\infty]$ , c'est à dire que  $\sup \emptyset = 0$ .

En effet soit  $\lambda \in ]0, +\infty[$  tel que  $\lambda > \alpha$ ; alors

$$\exists x_0 \in X : g(x_0) \geq 1, f(x_0) < \lambda$$

c'est à dire

$$\exists x_0 \in X : g\left(\frac{1}{\lambda} x_0\right) \geq \frac{1}{\lambda}, f\left(\frac{1}{\lambda} x_0\right) < 1 ;$$

donc, pour un tel  $\lambda$  (s'il en existe, c'est à dire si  $\alpha < +\infty$ , sans quoi (8.1) est trivial) on a  $\frac{1}{\lambda} \leq \beta'$ , c'est à dire  $\lambda \geq \frac{1}{\beta'}$ . Compte tenu des inégalités évidentes  $\alpha \leq \alpha'$  et  $\beta \geq \beta'$ , cela donne

$$(8.1) \quad \alpha' \geq \alpha \geq \frac{1}{\beta'} \geq \frac{1}{\beta} .$$

De même soit  $\mu \in ]0, +\infty[$  tel que  $\mu > \frac{1}{\beta}$  c'est à dire  $\frac{1}{\mu} < \beta$ ; alors

$$\exists x_1 \in X : f(x_1) \leq 1, g(x_1) > \frac{1}{\mu}$$

c'est à dire

$$\exists x_1 \in X : f(\mu x_1) \leq \mu, g(\mu x_1) > 1 ;$$

donc, pour un tel  $\mu$  (s'il en existe, c'est à dire si  $\beta > 0$  sans quoi (8.2) est trivial) on a  $\mu > \alpha'$ , donc

$$(8.2) \quad \frac{1}{\beta} \geq \alpha' .$$

Le rapprochement de (8.1) et (8.2) fournit la conclusion cherchée

REMARQUE.

Si  $f$  et  $g$  sont telles qu'on ait l'implication

$$f(x) < +\infty \Rightarrow g(x) < +\infty$$

(en particulier si  $g$  est partout finie)  $\alpha$  et  $\alpha'$  sont aussi égaux à  $\inf \{f(x) : g(x) = 1\}$ .

En effet ce dernier inf est essentiellement  $\geq \alpha$ . Pour montrer l'inégalité inverse supposons  $\lambda \in ]0, +\infty[$  tel que  $\lambda > \alpha$  (s'il n'existe pas de tel  $\lambda$ , le résultat est trivial); alors

$$\exists x_0 \in X : g(x_0) \geq 1, f(x_0) < \lambda .$$

En vertu de l'hypothèse  $g(x_0) \in [1, +\infty[$ , donc

$$\frac{1}{g(x_0)} \in ]0, 1]$$

et l'homogénéité donne

$$f\left(\frac{1}{g(x_0)} x_0\right) = \frac{1}{g(x_0)} f(x_0) < \lambda ,$$

tandis que

$$g\left(\frac{1}{g(x_0)} x_0\right) = 1 ,$$

donc

$$\lambda > \inf \{f(x) : g(x) = 1\}$$

ce qui prouve l'assertion.

Le contre exemple suivant montre que si  $f$  et  $g$  ne vérifient pas l'hypothèse faite, l'égalité annoncée peut ne pas avoir lieu : prendre  $g$  n'admettant que les valeurs  $0$  et  $+\infty$ . Alors  $\{x : g(x) = 1\} = \emptyset$  de sorte que

$$\inf \{f(x) : g(x) = 1\} = +\infty .$$

Par contre,  $\{x : g(x) \geq 1\} = \{x : g(x) = +\infty\}$  sera une partie de  $X$  (stable pour la multiplication par tout  $\lambda$ ) sur laquelle l'inf des valeurs de  $f$  peut être  $0$ .

Par un raisonnement semblable on établira :

Si et  $f$  et  $g$  sont telles qu'on ait l'implication

$$g(x) > 0 \Rightarrow f(x) > 0$$

(en particulier si  $f$  est partout  $> 0$ ),  $\beta$  et  $\beta'$  sont aussi égaux à  $\sup \{g(x) : f(x) = 1\}$ .

=====

BIBLIOGRAPHIE

CHENEY (E.W.), GOLDSTEIN (A.A.)

- [1] Tchebycheff approximation in locally convex spaces, Bull. Amer. Math. Soc., 68 (1962), 449-450.

CHOQUET (G.)

- [1] Les cônes convexes faiblement complets dans l'analyse, Proceedings of the International congress of Mathematicians, Stockholm 1962, p. 322.
- [2] Sur la meilleure approximation dans les espaces normés, Rev. Math. Pures Appl. 8 (1963), 541-542.

FAN (K.)

- [1] Minimax theorems, Proc. Nat. Acad. Sci., 39 (1953), 42-47.

FENCHEL (W.)

- [1] On conjugate convex functions, Canad. J. Math. 1 (1949), 73-77.
- [2] Convex cones, sets and functions, multilith. lecture notes, Princeton Univ. (1953).

GARKAVI (A.L.)

- [1] Théorèmes de dualité pour l'approximation au moyen d'éléments d'ensembles convexes (en russe), Uspekhi Mat. Nauk, 16 (1961), 141-145.

GOL'STEIN (E.G.)

- [1] On a general formulation of the problem of best approximation, Dokl. Akad. Nauk SSSR, 144 (1962), 21 = Soviet Math. Dokl. 3 (1962), 637-639.

HAVINSON (S. Ya.)

- [1] Approximation by elements of convex sets, Dokl. Akad. Nauk SSSR, 172 (1967) N°2 = Soviet Math. Dokl. 8 (1967) 98-100.

KNESER (H.)

- [1] Sur un théorème fondamental de la théorie des jeux, C.R. Acad. Sci. Paris, 234 (1952), 2418-2420.

KRIPKE (B.R.), ROCKAFELLAR (R.T.)

- [1] A necessary condition for the existence of best approximations, J. Math. Mech., 13 (1964), 1037-1038.

LAURENT (P.J.)

- [1] Théorèmes de caractérisation d'une meilleure approximation dans un espace normé et généralisation de l'algorithme de Rémès, Num. Math., 10 (1967) 190- 208.

LESCARRET (C.)

- [1] Sur la sous-différentiabilité d'une somme de fonctionnelles convexes semi-continues inférieurement, C.R. Acad. Sci. Paris, 262 (1966) 443-446.

MOREAU (J.J.)

- [1] Sur la fonction polaire d'une fonction semi-continue supérieurement, C.R. Acad. Sci. Paris, 258 (1964), 1128-1131.
- [2] Théorèmes "inf-sup", C.R. Acad. Sci. Paris, 258 (1964), 2720-2722.
- [3] Fonctionnelles convexes, Séminaire sur les Equations aux Dérivées Partielles, Collège de France, Paris (1966-67) (multigraphié 108 pages).

- [4] Un cas d'addition des sous-différentiels, Travaux du Séminaire d'Analyse Unilatérale, Vol. 2 (1969), exposé n° 3.

ROCKAFELLAR (R.T.)

- [1] Extensions of Fenchels' duality theorem for convex functions, Duke Math. J. 33, (1966) 81-90.

ROUMIEU (Ch.)

- [1] Sur quelques problèmes d'approximation, Faculté des Sciences de Montpellier, Séminaires de Mathématiques, 1966 (multigraphié 27p.).

RUBINSTEIN (G.S.)

- [1] Dual extremum problems, Dokl. Akad. Nauk SSSR, 152 (1963), 288 = Soviet Math. Dokl. 4 (1963), 1309-1312.

SINGER (I.)

- [1] Caractérisation des éléments de meilleure approximation dans un espace de Banach quelconque, Acta Sci. Math. Szeged, 17 (1956), 181-189.

=====