

ENSEMBLES DE TCHEBYCHEV VS ENSEMBLES CONVEXES : L'ÉTAT DE LA SITUATION VU VIA L'ANALYSE CONVEXE NON LISSE

JEAN-BAPTISTE HIRIART-URRUTY

RÉSUMÉ. Un ensemble fermé S d'un espace de Hilbert H est dit de Tchebychev si tout point de H admet un et un seul point à distance minimale dans S . Un tel ensemble est-il nécessairement convexe? Cette question, l'une des plus célèbres en approximation et analyse réelle, n'a pas reçu de réponse complète à ce jour, du moins lorsque H est de dimension infinie. L'objet du travail présent est précisément de faire le point sur ce problème, en privilégiant le point de vue que permet l'analyse convexe non lisse.

ABSTRACT. A closed set S of a Hilbert space H is said to be Tchebychev if every point in H has one and only one closest point in S . Is such a set necessarily convex? This question, one of the most famous in Approximation and Real Analysis, has not been completely answered as yet, at least when H is infinite dimensional. The aim of the present work is precisely to take stock of this question, emphasizing the viewpoint of Convex Nonsmooth Analysis.

1. Introduction. Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert réel, structuré en espace vectoriel normé grâce à la norme hilbertienne associée à $\langle \cdot, \cdot \rangle$ et que nous notons $\| \cdot \|$. Étant donné un ensemble *fermé* non vide S de H et x un point de H , désignons par $d_S(x)$ la distance hilbertienne de x à S , et par $P_S(x)$ l'ensemble des points de S à distance minimale de x , c'est-à-dire :

$$(1.1) \quad \begin{aligned} \forall x \in H, \quad d_S(x) &:= \inf \{ \|x - s\| \mid s \in S \}, \\ P_S(x) &:= \{ s \in S \mid d_S(x) = \|x - s\| \}. \end{aligned}$$

L'ensemble $P_S(x)$ peut être vide, contenir plusieurs éléments, etc. L'opération qui à $x \in H$ associe $P_S(x)$ définit donc une *multiapplication* de H dans S . Parmi les qualificatifs de S issus des propriétés de P_S , citons ceux de la définition suivante :

Reçu le 7 mai 1996 et, sous forme définitive, le 30 janvier 1997.

Version écrite d'un exposé au colloque « Set-valued Analysis and Nonsmooth Analysis » (Saint-Petersbourg, juin 1995).

Définition 1.1.

- S est appelé ensemble *proximal* lorsque $P_S(x) \neq \emptyset$ pour tout $x \in H$.
- S est dit *de Tchebychev* lorsque $P_S(x)$ est un singleton (c'est-à-dire réduit à un seul élément) pour tout $x \in H$.

Une des propriétés des convexes, essentielle en approximation et en optimisation, et que nous enseignons tous à nos étudiants est la suivante : *tout ensemble convexe (fermé) S de H est de Tchebychev*. La question que l'on peut se poser ensuite est : la réciproque est-elle vraie? Autrement dit : *un ensemble de Tchebychev est-il nécessairement convexe*? La réponse est connue comme étant *oui* lorsque H est de dimension finie, et ce depuis plus de soixante ans; c'est le fameux théorème de L. N. H. Bunt (dans sa thèse en 1934), habituellement et incorrectement attribué à T. S. Motzkin (1935)¹, sur lequel nous reviendrons plus loin en proposant une démonstration par les techniques d'analyse convexe moderne. Et si H est de dimension infinie? La question a été clairement posée par V. Klee il y a près de trente cinq ans (vers 1961) et cet auteur conjecturait même que la réponse était non : il pensait qu'il existait un espace de Hilbert de dimension infinie (éventuellement non séparable) contenant un ensemble de Tchebychev non convexe. Et bien qu'il y ait eu des progrès considérables depuis, la question reste ouverte à ce jour.

Nous n'avons pas la prétention d'y répondre ici mais seulement de faire un état de la situation, en adoptant un point de vue peu habituel chez les analystes qui ont abordé ce problème, celui de l'analyse convexe et/ou non lisse. Une fonction jouera un rôle-clé dans cette approche : la fonction-distance d_S (et son carré d_S^2) ou plutôt la fonction φ_S (notre fonction préférée) définie comme étant $\frac{1}{2}[\|\cdot\|^2 - d_S^2]$. Deux propriétés de ces fonctions seront essentielles à considérer : leur *convexité* et leur *différentiabilité* (dans un sens qu'il faudra préciser). Toute notre démarche tournera autour de ces deux propriétés et leurs liens avec le caractère convexe ou de Tchebychev de l'ensemble S considéré.

Avant d'aller plus loin, il est bon d'observer que le problème posé est vraiment un problème *hilbertien*, où toutes les données en hypothèses comptent :

- Si $\|\cdot\|$ n'est pas une norme hilbertienne, la question n'a pas lieu d'être posée, même si H est de dimension finie. Par exemple, si $\mathbb{R}^2$ est muni de la norme $\|(x_1, x_2)\| := \max(|x_1|, |x_2|)$ et si $S := \{(x_1, x_2) \in \mathbb{R}^2 \mid (x_2 \geq x_1 \geq 0) \text{ ou } (x_2 \geq 2x_1 \text{ et } x_1 \leq 0)\}$ tout point de $\mathbb{R}^2$ admet un et un seul point à distance minimale dans S , alors que S est loin d'être convexe (cf. figure 1).
- Si $\|\cdot\|$ est la norme euclidienne usuelle de $\mathbb{R}^2$ et si $S := \{(x_1, x_2) \in \mathbb{R}^2 \mid \|(x_1, x_2)\| \geq 1\}$, tout point *sauf un* (l'origine) admet un et un seul point à distance minimale dans S , mais S est violemment non convexe.
- La question peut être posée dans un espace préhilbertien, mais nous verrons que le caractère complet ou pas de l'espace a son rôle à jouer.

2. Le problème en dimension finie. On suppose dans ce paragraphe que H est de *dimension finie* ($(H, \langle \cdot, \cdot \rangle)$ est donc ce qu'on appelle un espace euclidien). S étant un fermé non vide de H , on note S^c son complémentaire dans H .

¹Nous remercions F. Deutsch (The Pennsylvania State University) de nous avoir apporté des précisions historiques à ce sujet; nous appellerons désormais ce résultat le théorème de Bunt.

Dans le cas où $f : \mathcal{O} \subset H \rightarrow \mathbb{R}$ est différentiable en un point x de l'ouvert $\mathcal{O}$, la notation $\nabla f(x)$ est utilisée pour le gradient de f en x ($\nabla f(x)$ est l'unique élément de H représentant la différentielle $Df(x)$ de f en x , via $\langle \cdot, \cdot \rangle$).

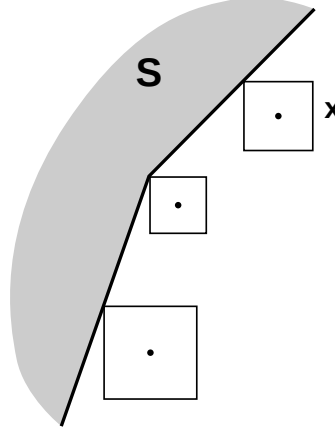


Figure 1

2.1. Différentiabilité de la fonction distance d_S vs celle de son carré. Première observation, résultant de la règle de différentiation de fonctions composées différentiables :

$$(2.1) \quad (d_S \text{ différentiable en } x) \implies (d_S^2 \text{ différentiable en } x).$$

Si x est intérieur à S , d_S et d_S^2 sont différentiables en x , avec $\nabla d_S(x) = \nabla d_S^2(x) = 0$. Si x n'appartient pas à S , c'est-à-dire si $d_S(x) > 0$, toujours en raison de la règle de différentiation invoquée plus haut, d_S est différentiable en x si et seulement si d_S^2 est différentiable en x . Ainsi, hormis en des points de la frontière $\text{Fr } S$ de S , les points de différentiabilité de d_S et de d_S^2 sont les mêmes :

$$(2.2) \quad \{x \notin \text{Fr } S \mid d_S \text{ est différentiable en } x\} = \{x \notin \text{Fr } S \mid d_S^2 \text{ est différentiable en } x\}.$$

La fonction d_S est Lipschitz sur H , de constante de Lipschitz 1. Par suite, si $x \in S$, on a $d_S^2(x+h) \leq \|h\|^2$, ce qui implique que d_S^2 est différentiable en x (c'est l'effet «adoucissant» de l'élévation au carré au voisinage de 0^+). En somme :

$$(2.3) \quad d_S^2 \text{ est toujours différentiable sur } S.$$

On aura à utiliser cette propriété par la conséquence suivante :

$$(2.4) \quad (d_S \text{ différentiable sur } S^c) \iff (d_S^2 \text{ différentiable sur tout } H).$$

La fonction d_S peut être différentiable en un point extérieur à S comme elle peut ne pas l'être (le test de différentiabilité sera précisément l'objet du §2.2). Et en un point de la frontière de S ? La fonction d_S peut être différentiable en certains points de $\text{Fr } S$ comme elle peut ne pas l'être en aucun des points de $\text{Fr } S$.

Exemple 2.1. La fonction d_S peut être différentiable en $x \in \text{Fr } S$ (auquel cas $\nabla d_S(x) = 0$ nécessairement). Soit par exemple $S := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \leq |x_1|^{1/2}\}$. Alors d_S est différentiable en $x = (0, 0)$.

Cette situation ne se produit pas s'il y a un «cône de sécurité» en x vers l'extérieur de S ; la figure 2 montre d'une manière suffisamment explicite ce qu'on entend par là.

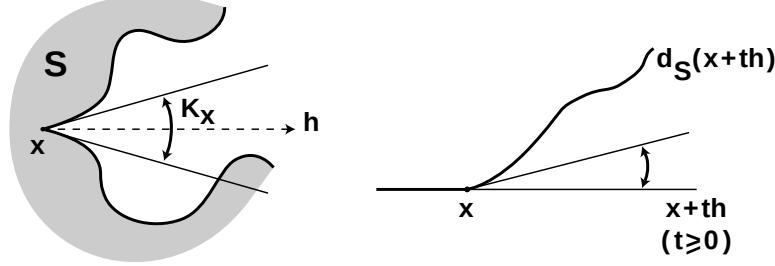


Figure 2

Signalons enfin le résultat suivant [33, p. 13] que l'on peut aussi dériver du théorème de Rademacher :

$$(2.5) \quad d_S(x+h) = o(\|h\|) \text{ pour presque tout } x \text{ de } S.$$

Exemple 2.2. La fonction d_S peut n'être différentiable en aucun point de $\text{Fr } S$. Si S est convexe, c'est carrément un demi-espace de sécurité K_x qu'il y a en tout $x \in \text{Fr } S$: la fonction d_S n'est différentiable en aucun point de $\text{Fr } S$.

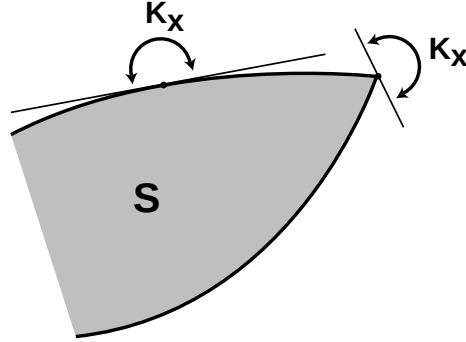


Figure 3

Remarque. Puisque H est de dimension finie et que d_S (resp. d_S^2) est Lipschitz (resp. localement Lipschitz) sur H , il n'y a pas lieu de distinguer entre la différentiabilité (usuelle) au sens de Fréchet et celles au sens de Gâteaux ou Hadamard. Toutefois les propriétés (2.1) à (2.4) n'utilisent pas la propriété de dimension finie de H et sont encore valables pour la différentiabilité sur un espace H de dimension infinie.

2.2. Test de différentiabilité de d_S en un point extérieur à S . Soit x extérieur à S ; comment reconnaître alors à l'aide de $P_S(x)$ si d_S (ou d_S^2) est différentiable en x ou non? La première chose à dire est la suivante : si d_S est différentiable en x , alors il n'y a qu'un point $\bar{x}$ dans S à distance minimale de x , c'est-à-dire $P_S(x) = \{\bar{x}\}$; auquel cas

$$(2.6) \quad \nabla d_S(x) = \frac{x - \bar{x}}{\|x - \bar{x}\|} \text{ (vecteur unitaire donc, dirigé par } x - \bar{x} \text{)} ;$$

$$(2.7) \quad \nabla \left(\frac{1}{2} d_S^2 \right) (x) = x - \bar{x}.$$

Cette propriété, fort ancienne et déjà remarquée par J. M. Bony [6, §5] et F. H. Clarke [10, Proposition 2.4], se démontre facilement et c'est d'ailleurs un excellent exercice de calcul différentiel. Plus intéressant et moins connu est que la réciproque est vraie : si $P_S(x) = \{\bar{x}\}$ (toujours avec $x \notin S$), alors d_S est différentiable en x (et, bien sûr, on a (2.6)). On peut trouver une démonstration de ceci dans la note de A. Pommellet [28, p. 378-379]. Il en résulte les équivalences suivantes :

$$(2.8) \quad \begin{aligned} (S \text{ est de Tchebychev}) &\iff (d_S \text{ est différentiable sur } S^c) \\ &\iff (d_S^2 \text{ est différentiable sur } H). \end{aligned}$$

2.3. Dérivée directionnelle de d_S . Étant donné un point x de H , on peut se poser la question de l'existence de la dérivée directionnelle (usuelle) de la fonction d_S en x et de son expression explicite. La réponse est comme suit : cette dérivée directionnelle

$$(2.9) \quad h \in H \mapsto d'_S(x, h) := \lim_{t \rightarrow 0^+} \frac{d_S(x + th) - d_S(x)}{t}$$

existe en tout point $x \notin S$ et vaut $\min \left\{ \left\langle \frac{x - s}{\|x - s\|}, h \right\rangle \mid s \in P_S(x) \right\}$.

Ainsi :

$$(2.10) \quad \text{Pour tout } x \notin S, \quad d'_S(x, h) = \min \left\{ \left\langle \frac{x - s}{\|x - s\|}, h \right\rangle \mid s \in P_S(x) \right\}.$$

$$(2.11) \quad \text{Pour tout } x \in H, \quad \left(\frac{1}{2} d_S^2 \right)' (x, h) = \min \{ \langle x - s, h \rangle \mid s \in P_S(x) \}.$$

On trouve trace de ce résultat dès les années trente, et on le retrouve régulièrement dans les articles ou livres : [39] et [19, Lemma 8.5.12] en sont des exemples. La démonstration n'en est pas très difficile, soit directement en utilisant les propriétés particulières de d_S et de P_S comme cela est fait classiquement, soit en appliquant des résultats de dérivation directionnelle du supremum d'une famille de fonctions différentiables tels que les contributions à l'optimisation moderne ont su dégager (voir [3, p. 56] par exemple).

Notons qu'en raison du caractère de Lipschitz de d_S , la convergence de

$$[d_S(x + th) - d_S(x)]/t - d'_S(x, h)$$

vers 0 est uniforme pour h dans un borné de H ; il n'y a donc pas lieu de faire une distinction entre les dérivées directionnelles au sens de Gâteaux, Hadamard ou Fréchet de d_S (ou de d_S^2).

Il est à présent clair que les formules (2.10) et (2.11) englobent les formules (2.6) et (2.7) établies en présence de différentiabilité en x , et que les équivalences de (2.8) s'ensuivent. F. H. Clarke pour lequel la fonction-distance fut un outil essentiel dans sa

démarche, avait en fait obtenu un résultat supérieur à (2.10) [10, Theorem 2.1]; il avait démontré que si $x \notin S$ la fonction $-d_S$ était *strictement tangentiellement convexe* (ou *régulière au sens de Clarke*) en x , c'est-à-dire

$$(2.12) \quad -d'_S(x, h) = (-d_S)^o(x, h) = \limsup_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{-d_S(x' + th) + d_S(x')}{t},$$

ainsi que l'expression (2.10) correspondante. Il s'ensuivait, entre autres, que si d_S était différentiable sur un ouvert $\mathcal{O} \subset S^c$ (ce que nous savons être équivalent à $P_S(x)$ est un singleton pour tout $x \in \mathcal{O}$), alors d_S était de classe C^1 sur $\mathcal{O}$.

Remarque. La fonction d_S ne sait pas distinguer $\text{Fr } S$ et $\overset{\circ}{S}$ puisqu'elle est nulle sur $\overset{\circ}{S} \cup \text{Fr } S$. Il existe une autre fonction, étudiée par l'auteur dans le cadre de l'analyse convexe et/ou non lisse [20, section II.B], qui parfois fait mieux l'affaire : si S est différent de H , on définit la *fonction-distance signée* (à S) Δ_S comme suit :

$$(2.13) \quad \text{Pour tout } x \in H, \Delta_S(x) = \begin{cases} d_S(x) & \text{si } x \notin S, \\ -d_{S^c}(x) & \text{si } x \in S. \end{cases}$$

Alors

$$\overset{\circ}{S} = \{x \mid \Delta_S(x) < 0\}, \text{Fr } S = \{x \mid \Delta_S(x) = 0\}, S^c = \{x \mid \Delta_S(x) > 0\}$$

et il se trouve que Δ_S est convexe si et seulement si S est convexe. La différentiabilité ou non de Δ_S s'étudie de la même manière que celle de d_S .

Signalons enfin que le caractère sous-harmonique de d_S (resp. surharmonique de Δ_S) a un lien très fort avec la convexité de S [1, 26].

2.4. Vers le théorème de Bunt. Les équivalences de (2.8) montrent clairement les voies de démonstration possibles pour la partie difficile du théorème de Bunt «(S est de Tchebychev) $\implies$ (S est convexe)», à savoir :

– démonstrations directes de

$$(2.14) \quad (S \text{ est de Tchebychev}) \implies (\text{Une propriété caractéristique de la convexité de } S);$$

– démonstrations partant de la différentiabilité de d_S^2

$$(2.15) \quad (d_S^2 \text{ est différentiable sur } H) \implies (S \text{ est convexe}).$$

À notre connaissance, toutes les démonstrations classiques consistent à démontrer l'implication (2.14), et sont de nature *géométrique* ([35, Theorem 7.8] par exemple, ou [28, p. 380-382] pour une preuve pédagogique et détaillée); ceci est confirmé par l'exposé historique de la question confectionné par F. Deutsch [14].

Nous proposons ci-dessous une autre démonstration du théorème de Bunt, de l'implication (2.15) en fait, que nous qualifierons de moderne car elle utilise les résultats et techniques de l'analyse convexe moderne. Cette démonstration a été publiée pour l'essentiel dans [21, Chapter X].

Posons

$$(2.16) \quad x \in H \mapsto f_S(x) := \begin{cases} 1/2\|x\|^2 & \text{si } x \in S, \\ +\infty & \text{sinon.} \end{cases}$$

Il est clair que f_S est semi-continue inférieurement sur H , qu'elle est 1-coercive sur H au sens où $\lim_{\|x\| \rightarrow +\infty} f_S(x)/\|x\| = +\infty$; de plus f_S est convexe si et seulement si S est convexe. Autrement dit, démontrer que S est convexe revient à démontrer que la fonction f_S est convexe. Le lemme suivant est crucial dans notre démarche car il fournit un test de convexité d'une fonction.

Lemme. Soit $f : H \rightarrow \mathbb{R} \cup \{+\infty\}$, non identiquement égale à $+\infty$, semi-continue inférieurement et 1-coercive sur H . Supposons que la transformée de Legendre-Fenchel f^* de f soit différentiable sur H . Alors f est nécessairement convexe.

Démonstration. Elle se fait en plusieurs étapes et on y utilise abondamment les résultats du §1.5 du chapitre X de [21].

Préliminaires. Les hypothèses faites sur f font que f^* est à valeurs finies sur H [21, Proposition 1.3.8]; ces mêmes hypothèses assurent que la convexifiée $\text{co}f$ de f s'obtient à partir de f par le procédé suivant [21, Proposition 1.5.4 (ii)] : pour tout x en lequel $\text{co}f$ est finie, il existe $x_1, \dots, x_{n+1}$ en lesquels f est finie et des coefficients $\alpha_1, \dots, \alpha_{n+1}$ positifs de somme égale à 1, tels que

$$(2.17) \quad x = \alpha_1 x_1 + \dots + \alpha_{n+1} x_{n+1} \text{ et } (\text{co}f)(x) = \alpha_1 f(x_1) + \dots + \alpha_{n+1} f(x_{n+1})$$

(n est ici la dimension de H).

Démontrer que f est convexe revient à démontrer que $\text{co}f = f$.

Calcul sous-différentiel. Désignons par $\partial g(u)$ le sous-différentiel (au sens de l'analyse convexe) de la fonction (non nécessairement convexe) g en u . Il est alors possible d'exprimer $\partial(\text{co}f)(x)$ à l'aide des $\partial f(x_i)$, où les x_i sont des points parmi ceux mis en évidence en (2.17). De fait [21, Theorem 1.5.6], si les points x_i et les coefficients α_i sont choisis de manière à avoir (2.17), on a :

$$(2.18) \quad \partial(\text{co}f)(x) = \bigcap_{\{i|\alpha_i>0\}} \partial f(x_i).$$

Prise en compte de la différentiabilité de f^* . Les hypothèses de semi-continuité inférieure et de 1-coercivité de f font que $\text{co}f$ est elle aussi semi-continue inférieurement [21, Theorem 1.5.4 (i)]. Par suite : $(x \in \partial f^*(x^*)) \iff (x^* \in \partial(\text{co}f)(x))$.

Supposons que x s'exprime comme cela est requis en (2.17) à l'aide de points x_i différents de x et des coefficients $\alpha_i > 0$. On aurait alors d'après l'équivalence qui précède et (2.18) :

$$(x = \nabla f^*(x^*)) \iff (x^* \in \partial(\text{co}f)(x) = \bigcap_{\{i|\alpha_i>0\}} \partial f(x_i)),$$

d'où $x_i \in \partial f^*(x^*)$, ce qui contredit la différentiabilité de f^* . Les x_i en question sont donc tous égaux à x , ce qui fait que $(\text{co}f)(x) = f(x)$. $\square$

Remarque. Comme $(cof)(x_i) = f(x_i)$ en tout point x_i de la famille considérée en (2.17), la démonstration du lemme montre en fait un peu plus : cof est strictement convexe sur toute partie convexe contenue dans $\{x \in H \mid \partial(cof)(x) \neq \emptyset\}$; la fonction cof est donc ce que R. T. Rockafellar appelle « essentiellement strictement convexe » [29, p. 253].

Revenons à notre fonction f_S . Quelle est sa transformée de Legendre-Fenchel ? Elle est très facile à calculer :

$$\begin{aligned} \forall x^* \in H, f_S^*(x^*) &= \sup_{s \in S} \left\{ \langle x^*, s \rangle - \frac{1}{2} \|s\|^2 \right\} \\ &= \frac{1}{2} \|x^*\|^2 - \frac{1}{2} \sup_{s \in S} \|x^* - s\|^2 \\ &= \frac{1}{2} \left[\|x^*\|^2 - d_S^2(x^*) \right]. \end{aligned}$$

On reviendra plus longuement au paragraphe suivant sur cette fonction $x^* \in H \mapsto \varphi_S(x^*) := \frac{1}{2} [\|x^*\|^2 - d_S^2(x^*)]$; mais il est immédiat à présent que

$$(2.19) \quad (d_S^2 \text{ différentiable sur } H) \iff (\varphi_S \text{ est différentiable sur } H).$$

Grâce au lemme de test de convexité, on voit donc que la différentiabilité de d_S^2 (et donc celle de φ_S) implique la convexité de f_S .

Signalons pour être complet qu'il existe une démonstration du théorème de Bunt basée sur un théorème de point fixe et attribuée à L. P. Vlasov (1965); voir §2 de l'Appendice I de [30] ou [34, p. 63].

3. Les difficultés du problème en dimension infinie. Nous revenons dans ce paragraphe à la situation générale d'un fermé non vide S contenu dans un espace de Hilbert $(H, \langle \cdot, \cdot \rangle)$ de dimension infinie.

3.1. Faits généraux concernant la multiapplication de projection P_S . La définition même de $P_S(x)$ et le caractère lipschitzien de la fonction d_S sur H conduisent facilement aux propriétés suivantes :

- (3.1) P_S est à valeurs fermées bornées.
- (3.2) Le graphe de P_S , c'est-à-dire $\{(x, y) \in H \times H \mid y \in P_S(x)\}$ est fermé.
- (3.3) P_S est localement bornée, au sens où elle transforme les bornés en bornés.

Plus intéressantes et moins immédiates :

- (3.4) P_S est monotone, et même cycliquement monotone; par contre elle ne saurait être monotone maximale (à moins que S ne soit convexe, auquel cas P_S est univoque).
- (3.5) $P_S(x) = \{p_S(x)\}$ pour « beaucoup » de points x de H (le « beaucoup » pouvant prendre plusieurs sens, l'un d'entre eux étant « sur une partie dense de H »).

Ces dernières propriétés qualitatives figurent parmi celles de *sous-gradients de fonctions convexes*. En fait

- (3.6) Il existe une fonction convexe continue $\varphi_S : H \rightarrow \mathbb{R}$ telle que

$$P_S(x) \subset \partial \varphi_S(x) \text{ pour tout } x \in H.$$

Ainsi toute propriété du sous-différentiel d'une fonction convexe se transmet via l'inclusion ci-dessus à la multiapplication P_S , et c'est d'ailleurs comme cela qu'on voit le plus rapidement (3.4) et (3.5). Comme fonction φ_S faisant l'affaire, il y a la fonction à laquelle nous avons fait allusion à la fin de la section 2.4 et qui est définie comme suit :

$$(3.7) \quad \forall x \in H, \quad \varphi_S(x) := \frac{1}{2} \left[\|x\|^2 - d_S^2(x) \right].$$

Nous appellerons φ_S *fonction d'Asplund associée à S* , en hommage à celui qui, sans doute le premier, a utilisé de manière profonde cette fonction [2]. Avouons que c'est notre fonction favorite dans le contexte des questions évoquées dans ce travail. La fonction φ_S (partout finie et continue) possède la propriété remarquable d'être *toujours convexe*, quel que soit S donc. La façon la plus simple de le voir est d'observer que

$$\varphi_S(x) = \sup_{s \in S} \left\{ \langle x, s \rangle - \frac{1}{2} \|s\|^2 \right\} \text{ pour tout } x \in H,$$

φ_S est le supremum d'une famille de fonctions affines continues, en fait c'est la transformée de Legendre-Fenchel de la fonction (non nécessairement convexe)

$$f_S : x \in H \mapsto f_S(x) = \begin{cases} \frac{1}{2} \|x\|^2 & \text{si } x \in S, \\ +\infty & \text{sinon} \end{cases}$$

déjà introduite dans la section 2.4. La définition même du sous-différentiel $\partial\varphi_S$ ou sa caractérisation à l'aide de $(f_S)^* = \varphi_S$ permettent de voir facilement l'inclusion $P_S(x) \subset \partial\varphi_S(x)$; donc :

$$(3.8) \quad \overline{\text{co}}P_S(x) \subset \partial\varphi_S(x) \text{ pour tout } x \in H.$$

Rappelons que lorsque S est convexe, φ_S est (Fréchet-)différentiable sur H avec $\nabla\varphi_S(x) = p_S(x)$ pour tout $x \in H$; la fonction φ_S n'est autre que la primitive de l'application p_S sur H (primitive définie à une constante additive près).

3.2. Convexité vs propriété de Tchebychev : revue des principaux résultats connus.

Supposons S de Tchebychev. Tous les résultats connus à ce jour tendant à prouver que S est convexe sont du type suivant :

$$\left(\begin{array}{l} S \text{ est Tchebychev} \\ + \text{ condition } C_i \end{array} \right) \implies (S \text{ est convexe}).$$

Les différentes conditions C_i mises en évidence dans la littérature partagent toutes les propriétés suivantes :

- elles sont satisfaites lorsque S est convexe (de sorte que l'implication ci-dessus peut être complétée en une équivalence);
- elles sont automatiquement vérifiées quand H est de dimension finie.

Les mathématiciens se sont évertués au cours du temps à rendre ces conditions C_i les plus faibles possibles; des résultats très fins et ingénieux ont ainsi été proposés,

une description historique figure dans [14, section 2]. Nous citons ci-dessus quelques conditions C_i parmi les plus palpables.

(C_1) S est faiblement fermé (V. Klee, 1961).

C'est sans doute l'un des résultats les plus spectaculaires dans le domaine; il est revu via des méthodes d'analyse non lisse par F. H. Clarke *et al.* [13, section 4].

(C_2) L'application p_S est continue (L. P. Vlasov, 1967; E. Asplund, 1969).

Ici la continuité de $p_S : H \rightarrow H$ peut être entendue avec l'espace d'arrivée muni de la topologie forte ou muni de la topologie faible; les deux impliquent la convexité de S .

Il suffit en fait d'assurer une continuité «radiale» de p_S , c'est-à-dire : pour tout $x \in H$, la restriction de p_S à la demi-droite issue de x et dirigée par $x - p_S(x)$ est continue en x .

(C_3) S est «approximativement compact» (N. Efimov et S. Stechkin, 1961), au sens suivant : pour tout $x \in H$ et toute suite (s_n) de S telle que $\|x - s_n\| \rightarrow d_S(x)$ ((s_n) est une suite minimisante dans le problème de minimisation définissant $d_S(x)$), il existe une sous-suite convergente de (s_n) (la limite de cette sous-suite est alors $p_S(x)$ nécessairement).

(C_4) L'application p_S est monotone maximale.

Dans ce cas p_S est continue de H fort dans H faible, et on rejoint la condition (C_2); pour voir rapidement cette propriété de continuité, observons que sous la condition (C_4) l'inclusion (3.8) devient une égalité, la fonction (convexe continue) φ_S est différentiable avec $\nabla \varphi_S(x) = p_S(x)$ pour tout $x \in H$; la continuité annoncée s'ensuit [16, Proposition 2.5].

(C_5) L'ensemble des points de discontinuité de p_S est dénombrable (V. S. Balaganskii, 1982).

L'ensemble des points de discontinuité de l'application p_S est analysé par U. Westphal et J. Frerking [38]; il résulte de cette étude, entre autres, que si S (de Tchebychev) se trouve ne pas être convexe, l'ensemble des points de discontinuité de p_S contient une courbe lipschitzienne non réduite à un point. Cela implique le résultat de V. S. Balaganskii.

En 1987, G. G. Johnson [23] proposa un *ensemble de Tchebychev S non convexe* dans un espace *préhilbertien*; sa construction, loin d'être simple mais s'appuyant sur une démarche classique semble-t-il dans un tel contexte, comportait des imprécisions ou erreurs qui ont été corrigées récemment par M. Jiang [22]. Cette contribution au problème «Tchebychev vs convexe» soulève la question : quel est le rôle du caractère complet d'un espace de Hilbert? En tout cas, considérer l'espace complété du préhilbertien utilisé par G. G. Johnson ne permet pas de conclure. Les questions de différentiabilité de la fonction φ_S éclairent partiellement ce point.

3.3. Les difficultés propres à la dimension infinie. Ici nous analysons les difficultés empêchant de conclure en dimension infinie à la lumière des propriétés (de différentiabilité notamment) de la fonction d'Asplund φ_S associée à S . Lorsque H est de dimension finie, on sait [21, Proposition 3.2.2 du chapitre X] que

$$\partial \varphi_S(x) = \text{co}P_S(x) \text{ pour tout } x \in H,$$

ce qui permet de relier directement la propriété « S est de Tchebychev» à « φ_S est

différentiable». Qu'en est-il de la situation en dimension infinie? L'inclusion (3.8) pourrait-elle être stricte dans ce cas? Cette question, posée explicitement par H. Berens dans [5] a reçu une réponse par G. Godini [18] :

Dans tout espace de Hilbert H de dimension infinie, il existe un ensemble fermé S (proximal même) tel que

$$\overline{\text{co}}P_S(x) \subsetneq \partial\varphi_S(x) \text{ pour un certain } x \in H.$$

Ce résultat, fort étonnant par sa précision, repousse donc l'accès à la résolution de notre problème. Si S est de Tchebychev, on peut simplement affirmer que p_S fournit une sélection de sous-gradients de φ_S .

Comme en dimension finie (cf. 2.4) la différentiabilité de φ_S ou, ce qui revient au même, celle de d_S^2 conduit à la convexité de S . Le théorème ci-dessous rassemble les résultats les plus pertinents à cet égard.

Théorème 3.3.1 [2, 15, 32]. *Les assertions suivantes sont équivalentes :*

- (i) φ_S est Fréchet-différentiable sur H ;
- (ii) d_S^2 est Fréchet-différentiable sur H ;
- (iii) d_S est Fréchet-différentiable sur S^c ;
- (iv) S est convexe.

Si S est de Tchebychev, on peut ajouter l'équivalence avec

- (v) d_S^2 est Gâteaux-différentiable sur H ;
- (vi) d_S est Gâteaux-différentiable sur S^c .

À ce stade, on peut se demander quelle est l'expression exacte de $\partial\varphi_S(x)$ en termes non pas d'éléments de $P_S(x)$, ce que nous savons être insuffisant, mais d'éléments de S à distance «presque minimale» de x . Cette expression est connue (voir [5, p. 5] par exemple), nous la rappelons ci-dessous. Pour tout $\varepsilon > 0$, on définit

$$P_S^\varepsilon(x) := \{s \in S \mid \|x - s\| < d_S(x) + \varepsilon\}.$$

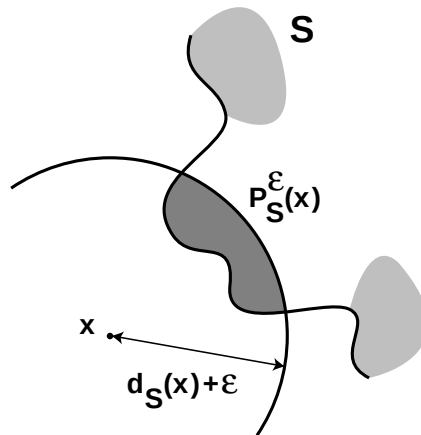


Figure 4

Alors on a :

$$(3.9) \quad \partial\varphi_S(x) = \bigcap_{\varepsilon>0} \overline{\text{co}}P_S^\varepsilon(x) \text{ pour tout } x \in H.$$

Comme conséquence, on observe que des formulations des dérivées directionnelles de $\frac{1}{2}d_S^2$ en $x \in H$ (ou de d_S en $x \in S^c$) aussi simples que celles de (2.10) et (2.11) sont sans espoir. Toutefois la question mérite d'être posée pour le cas où S est de Tchebychev et rejoint la question qui, au point où nous nous sommes rendus, devient centrale :

(Q) « S est de Tchebychev » implique-t-il « d_S^2 est Gâteaux-différentiable sur H » ?

Sachant que p_S est une sélection du sous-différentiel $\partial\varphi_S$ de la fonction convexe φ_S , il suffit de peu pour qu'il en soit ainsi : par exemple, que p_S (ou une autre sélection de $\partial\varphi_S$) soit continue de H fort dans H faible.

Rappelons que l'ensemble des points de Fréchet-différentiabilité de d_S^2 est connu (voir [2, 38]) :

$$(3.10) \quad (d_S^2 \text{ Fréchet-différentiable en } x) \iff (P_S(x) \text{ est un singleton et } P_S \text{ est semi-continue supérieurement en } x).$$

(La semi-continuité supérieure en x (où $P_S(x) = \{p_S(x)\}$) signifie : pour tout voisinage W de $p_S(x)$, on peut trouver un voisinage V de x tel que, pour tout x' de V , $P_S(x') \subset W$. Propriété toujours satisfaite lorsque H est de dimension finie.)

À défaut de (Gâteaux ou Fréchet)-différentiabilité, on peut se tourner vers des *différentiabilités généralisées* de d_S^2 (ou d_S). Deux d'entre elles ont été particulièrement étudiées, et les propriétés établies à leur sujet contribuent à éclairer le problème qui nous préoccupe. La première conduit au *sous-différentiel (ou gradient) généralisé au sens de Clarke* et la deuxième au *sous-différentiel proximal*. Voici les définitions : étant donné une fonction $f : H \rightarrow \mathbb{R}$ localement lipschitzienne sur H , le sous-différentiel généralisé (au sens de Clarke) de f en x est

$$(3.11) \quad \partial^{\text{cl}} f(x) = \{x^* \in H \mid \langle x^*, d \rangle \leq f^o(x, d) \text{ pour tout } d \in H\},$$

où

$$f^o(x, \cdot) : d \in H \mapsto f^o(x, d) := \limsup_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{f(x' + td) - f(x')}{t};$$

le sous-différentiel proximal de f en x est l'ensemble noté $\partial_{\text{prox}} f(x)$ des $x^* \in H$ vérifiant la propriété de « minoration quadratique » suivante : il existe α et η positifs tels que

$$(3.12) \quad f(x') \geq f(x) + \langle x^*, x' - x \rangle - \alpha \|x' - x\|^2$$

pour tout x' vérifiant $\|x' - x\| \leq \eta$. Les deux objets, $\partial^{\text{cl}} f(x)$ et $\partial_{\text{prox}} f(x)$, sont le « plus gros » et « le plus petit » des concepts associés à f en x et généralisant peu ou prou la notion de gradient; ils se confondent tous les deux en le sous-différentiel (usuel)

lorsque f est convexe. Les règles de calcul connues sur ces deux sous-différentiels [11, 12] donnent lieu aux expressions suivantes :

$$(3.13) \quad \partial^{\text{cl}}(d_S^2)(x) = 2[x - \partial\varphi_S(x)] \text{ pour tout } x \in H,$$

$$(3.14) \quad \partial^{\text{cl}}d_S(x) = \frac{x - \partial\varphi_S(x)}{d_S(x)} \text{ pour tout } x \in S^c.$$

$$(3.15) \quad \partial_{\text{prox}}(-d_S^2)(x) = 2[\partial\varphi_S(x) - x] \text{ pour tout } x \in H,$$

d'où

$$(3.16) \quad \partial_{\text{prox}}(-d_S^2)(x) = \partial^{\text{cl}}(-d_S^2)(x) \text{ pour tout } x \in H.$$

À notre connaissance il n'y a pas de relation simple reliant $\partial_{\text{prox}}(-d_S^2)(x)$ à $\partial_{\text{prox}}d_S(x)$ ou à $\partial_{\text{prox}}(-d_S)(x)$, simple au sens où il n'est fait appel qu'au point x considéré et non aux sous-différentiels en des points voisins de x . Signalons aussi que, parce que nous nous adressons à des fonctions particulières comme la fonction-distance, le sous-différentiel généralisé au sens de Clarke et des variantes de celui-ci comme le sous-différentiel généralisé au sens de Michel-Penot coïncident pour d_S en $x \in S^c$ [7].

Dans [13, Theorem 2.1], les auteurs prouvent l'implication suivante :

$$(3.17) \quad (\partial_{\text{prox}}d_S(x) \neq \emptyset \text{ pour tout } x \in S^c) \implies (S \text{ est de Tchebychev}).$$

On ne sait si l'implication réciproque est vraie; si c'était le cas, cela permettrait de conclure à la convexité des ensembles de Tchebychev via le théorème 3.3.1 car [13, Theorem 3.1] :

$$(3.18) \quad \left(\begin{array}{l} S \text{ est proximal et} \\ \partial_{\text{prox}}d_S \neq \emptyset \text{ pour tout } x \in S^c \end{array} \right) \implies (d_S \text{ est différentiable sur } S^c).$$

Remonter de « S est de Tchebychev » à « S est convexe » a pu être fait par ces techniques dans [13, section 4] au moins dans le cas où S est faiblement fermé, ce qui permet de retrouver le résultat de V. Klee (cf. §3.2). Il n'en reste pas moins que la question centrale revient à celle posée en **(Q)**, à savoir la Gâteaux-différentiabilité de $-d_S$ sur S^c . Observons à ce sujet qu'en raison du fait que $\|\cdot\|^2 - d_S^2$ est une fonction convexe continue sur H (c'est $2\varphi_S$ exactement), d_S est toujours directionnellement différentiable sur S^c avec

$$(3.19) \quad -d'_S(x, \cdot) = \frac{\varphi'_S(x, \cdot) - \langle x, \cdot \rangle}{d_S(x)} \quad \text{pour tout } x \in S^c,$$

ce qui permet aussi de voir immédiatement que $-d_S$ est toujours strictement tangentielllement convexe (ou régulière au sens de Clarke) sur S^c (chose déjà vue lorsque H est de dimension finie, cf. (2.12)). En définitive, **(Q)** se réduit à :

$$(\mathbf{Q}') \quad (S \text{ est de Tchebychev}) \stackrel{?}{\implies} (d'_S(x, \cdot) \text{ est linéaire pour tout } x \in S^c).$$

Pour avoir la Gâteaux-différentiabilité en $x \in S^c$ il a même été observé qu'il suffisait en fait de s'assurer de la Gâteaux-différentiabilité directionnelle de d_S dans la direction $x - p_S(x)$ (cf. [7]).

4. Conclusion. Après ce panorama, on est partagé entre deux points de vue opposés :

- Les progrès successifs enregistrés par les mathématiciens dans l'affaiblissement des hypothèses permettant d'assurer que « S est de Tchebychev» implique « S est convexe» laissent penser que cette implication est vraie; la «rigidité» de la structure hilbertienne conduit aussi à penser que la situation ne doit guère différer du cas euclidien.
- La construction de G. G. Johnson va au contraire dans le sens de la conjecture de V. Klee et, selon certains, a clos le problème. Le nombre infini de «degrés de liberté» dans un espace hilbertien H de dimension infinie peut laisser penser qu'on peut effectivement construire un ensemble de Tchebychev dans H qui ne soit pas convexe.

La question reste donc ouverte et il nous semble que l'approche par la différentiabilité de fonctions comme la fonction-distance serait un moyen d'y répondre.

«Il n'y a pas de problèmes résolus, il n'y a que des problèmes plus ou moins résolus»
H. Poincaré.

English extended abstract. Let H be a (real) Hilbert space and S a nonempty closed subset of H ; for all $x \in H$ we define $P_S(x)$ as the set of points in S closest to x for the hilbertian distance (see definitions in (1.1)). The set S is called *Tchebychev* when $P_S(x)$ is a singleton for all $x \in H$. A famous question, still not answered completely, is the following one: *is a Tchebychev set necessarily convex?*

The problem in a finite-dimensional context (H is finite-dimensional). The question above was answered (positively) by L. Bunt (thesis in 1934), and the corresponding result is usually (but incorrectly) attributed to T. S. Motzkin (1935). Here we look at this problem mainly from the viewpoint of the differentiability of the distance-function d_S (or of its square d_S^2). Examples of relationships we get:

$$\begin{aligned} (S \text{ is Tchebychev}) &\iff (d_S \text{ is differentiable on the complement of } S) \\ &\iff (d_S^2 \text{ is differentiable on the whole } H). \end{aligned}$$

By using techniques and results from modern Convex Analysis (subdifferential calculus, Legendre-Fenchel transformation) we provide a proof of the convexity of S when d_S^2 is differentiable on H .

The difficulties in an infinite-dimensional context. We first review the main additional assumptions known to ensure that a Tchebychev set is indeed convex. We do that by emphasizing again the role of the differentiability of the distance-function (Theorem 3.3.1). A key-role in that respect is played by the so-called Asplund function φ_S associated with S (cf. (3.7)). Such a function φ_S is always convex and the link between $P_S(x)$ and the subdifferential (in the sense of convex analysis) $\partial\varphi_S(x)$ of φ_S at $x \in H$ (see (3.8)) is the cornerstone of the (specific) difficulties encountered when the underlying Hilbert space is not finite-dimensional. Furthermore, since d_S as well as d_S^2 are locally Lipschitz functions, it is natural to ask whether considering their *generalized subdifferentials* (or *gradients*) in Clarke's sense is helpful to make some progress toward d_S tackling the problem. We provide some basic results in that direction

(see formulas (3.13) to (3.16)) and pose in a clear-cut manner the two central questions involving the differentiability of the distance-function, namely **(Q)** and **(Q')**, which (if answered) would allow us to answer completely the question on the convexity of a Tchebychev set.

BIBLIOGRAPHIE

1. D. H. Armitage et Ü. Kuran, *The convexity of a domain and the superharmonicity of the signed distance function*, Proc. Amer. Math. Soc. **93** (1985), 598–600.
2. E. Asplund, *Čebyšev sets in Hilbert space*, Trans. Amer. Math. Soc. **144** (1969), 235–240.
3. A. Auslender, *Optimisation. Méthodes numériques*, Masson, Paris, 1976.
4. V. S. Balaganskii, *Approximation properties of sets in Hilbert space*, Mat. Zametki **31** (1982), 785–800. (Russian)
5. H. Berens, *Best approximation in Hilbert space*, Approximation Theory III (Proc. Conf., Univ. Texas, Austin, Tex., 1980) (E. W. Cheney, ed.), Academic Press, New York, 1980, pp. 1–20.
6. J. M. Bony, *Quasi-analyticité et unicité du problème de Cauchy pour les solutions d'équations aux dérivées partielles*, Exposé n° 10 du Séminaire Goulaouic-Schwartz (1970–1971).
7. J. M. Borwein, S. P. Fitzpatrick et J. R. Giles, *The differentiability of real functions on normed linear space using generalized subgradients*, J. Math. Anal. Appl. **128** (1987), 512–534.
8. A. L. Brown, *An approximation-theoretic characterisation of inner product spaces*, Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988), 16–23.
9. J. V. Burke, M. C. Ferris et M. Qian, *On the Clarke subdifferential of the distance function of a closed set*, J. Math. Anal. Appl. **166** (1992), 199–213.
10. F. H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.
11. F. H. Clarke, *Optimization and nonsmooth analysis*, Canadian Mathematical Society series of Monographs and Advances Texts (John Wiley & Sons, ed.), Wiley-Interscience, New York, 1983; *Optimization and nonsmooth analysis. Second edition*, Classics in Applied Mathematics, vol. 5, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.
12. F. H. Clarke, *Analyse non lisse et optimisation*, Cours de l'école doctorale de mathématiques (1993–1994), Université Paul Sabatier de Toulouse.
13. F. H. Clarke, R. J. Stern et P. R. Wolenski, *Proximal smoothness and the lower- C^2 property*, J. Convex Anal. **2** (1995), 117–144.
14. F. Deutsch, *The convexity of Chebyshev sets in Hilbert space*, Topics in polynomials of one and several variables and their applications (Th. M. Rassias, H. M. Srivastava and A. Yanushauskas, eds.), World Scientific, River Edge, NJ, 1993, pp. 143–150.
15. S. Fitzpatrick, *Metric projections and the differentiability of distance functions*, Bull. Austral. Math. Soc. **22** (1980), 291–312.
16. J. R. Giles, *Convex analysis with application in the differentiation of convex functions*, Research Notes in Mathematics, vol. 58, Pitman, Boston, Mass., 1982.
17. J. R. Giles, *A distance function property implying differentiability*, Bull. Austral. Math. Soc. **39** (1989), 59–70.
18. G. Godini, *On a problem of H. Berens*, Math. Ann. **263** (1983), 279–281.
19. L. Hormander, *The analysis of linear partial differential operators I*, Distribution theory and Fourier analysis. Second edition, Springer-Verlag, New York-Berlin, 1989.
20. J.-B. Hiriart-Urruty, *New concepts in nondifferentiable programming*, Mémoire n° 60 du Bulletin de la Société Mathématique de France (1979), 57–85.

21. J.-B. Hiriart-Urruty et C. Lemaréchal, *Convex analysis and minimization algorithms II. Advanced theory and bundle methods*, Grundlehren der mathematischen Wissenschaften, vol. 306, Springer-Verlag, New York-Berlin, 1993.
22. M. Jiang, *On Johnson's example of a nonconvex Chebyshev set*, J. Approx. Theory **74** (1993), 152–158.
23. G. G. Johnson, *A nonconvex set which has the unique nearest point property*, J. Approx. Theory **51** (1987), 289–332.
24. V. Klee, *Convexity of Chebyshev sets*, Math. Annalen **142** (1961), 292–304.
25. T. D. Narang, *Convexity of Chebyshev sets*, Nieuw Archief Voor Wiskunde XXV (1977), 377–402.
26. M. J. Parker, *Convex sets and subharmonicity of the distance function*, Proc. Amer. Math. Soc. **103** (1988), 503–506.
27. R. R. Phelps, *Convex functions, monotone operators and differentiability. Second edition*, Lectures Notes in Mathematics, vol. 1364, Springer-Verlag, New York-Berlin, 1993.
28. A. Pommellet, *Différentiabilité de la fonction distance à une partie*, Revue de Mathématiques Spéciales n° 7 (1992-1993), 377–382.
29. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
30. I. Singer, *Best approximation in normed linear spaces by elements of linear subspaces*, Die Grundlehren der mathematischen Wissenschaften Band 171, Springer-Verlag, New York-Berlin, 1970.
31. V. Soloviov, *Duality for nonconvex optimization and its applications*, Anal. Math. **19** (1993), 297–315.
32. V. Soloviov, *Characterizations of convexity in terms of smoothness*, Prépublication, Department of Probability Theory, Moscow Aviation Institute, 1995.
33. E. Stein, *Singular integrals and differential properties of functions*, Princeton University Press, Princeton, NJ, 1970.
34. V. M. Tikhomirov, *Convex analysis*, Analysis II (R.V. Gamkrelidze, ed.), Springer-Verlag, New York-Berlin, 1990.
35. F. A. Valentine, *Convex sets*, McGraw-Hill, New York, 1964.
36. L. P. Vlasov, *Chebyshev sets and approximately convex sets*, Mat. Zametki **2** (1967), 191–200.
37. L. P. Vlasov, *Approximate properties of sets in normed linear spaces*, Russian Mathematical Surveys **28** (1973), 1–66.
38. U. Westphal et J. Frerking, *On a property of metric projections onto closed subsets of Hilbert spaces*, Proc. Amer. Math. Soc. **105** (1989), 644–651.
39. L. Zajicek, *Differentiability of the distance function and points of multi-valuedness of the metric projection in Banach space*, Czechoslovak Math. J. **33** (1983), 292–308.

J.-B. HIRIART-URRUTY

LABORATOIRE APPROXIMATION ET OPTIMISATION

UFR MATHÉMATIQUES, INFORMATIQUE, GESTION

UNIVERSITÉ PAUL SABATIER (TOULOUSE III)

118, ROUTE DE NARBONNE

31062 TOULOUSE CEDEX

FRANCE

E-MAIL : jbhuc@cict.fr