BULLETIN DE LA S. M. F.

J.J. MOREAU

# Proximité et dualité dans un espace hilbertien

*Bulletin de la S. M. F.*, tome 93 (1965), p. 273-299
<http://www.numdam.org/item?id=BSMF_1965__93__273_0>

*Bull. Soc. math. France,*
93, 1965, p. 273 à 299.

# PROXIMITÉ ET DUALITÉ DANS UN ESPACE HILBERTIEN ;

PAR

JEAN-JACQUES MOREAU
(Montpellier).

## 1. Introduction.

1.*a*. — Soit $H$ un espace hilbertien réel (ce qui suit peut évidemment concerner aussi un espace hilbertien complexe, par la structure hilbertienne réelle sous-jacente). Soit $C$ une partie convexe fermée non vide de $H$ et soit $z \in H$; on sait que l'application $u \to \| z - u \|$ de $C$ dans $\mathbf{R}$ atteint sa borne inférieure en un unique point $x$ de $C$, usuellement appelé *projection* de $z$ sur $C$ (notons $x = \mathrm{proj}_C\, z$).

En particulier, si $C$ est un sous-espace vectoriel fermé de $H$, on retrouve la *projection orthogonale* sur un tel sous-espace. La décomposition classique de $H$ en somme directe de deux sous-espaces supplémentaires orthogonaux se généralise utilement dans la propriété suivante, déjà publiée dans [15] et qui apparaîtra ici comme un corollaire de notre proposition 4.*a* : Soient, dans $H$ deux *cônes convexes fermés*, $P$ et $Q$ *mutuellement polaires*, c'est-à-dire que

$$P = \{ x \in H : (x \mid y) \leq 0 \text{ pour tout } y \in Q \}$$

(relation symétrique entre $P$ et $Q$). Alors tout $z \in H$ est égal à la somme de $x = \mathrm{proj}_P z$ et de $y = \mathrm{proj}_Q z$; on a $(x \mid y) = 0$ et c'est là *l'unique décomposition* de $z$ en somme d'un élément de $P$ et d'un élément de $Q$ *orthogonaux*.

1.*b*. — Les besoins de la mécanique, plus précisément la *dynamique des systèmes matériels à liaisons unilatérales* (*cf.* [16], [21], [27]) et aussi la *statique* de tels systèmes (milieu élastique de compressibilité limitée de W. PRAGER [29], [30], [31]) nous ont conduit à l'extension suivante de l'idée de projection sur un convexe fermé :

La notation $\Gamma_0(H)$ désignera l'ensemble des fonctions à valeurs dans $]-\infty, +\infty]$, partout définies sur $H$, convexes, semi-continues inférieurement, non partout égales à $+\infty$. Soit $f \in \Gamma_0(H)$ et soit $z$ quelconque dans $H$; on montrera que la fonction numérique $u \to \frac{1}{2} \| z - u \|^2 + f(u)$ atteint sa borne inférieure en un unique point $x$ noté $\text{prox}_f z$. On retrouve la projection sur un ensemble convexe fermé $C$ en prenant pour $f$ la *fonction indicatrice* de cet ensemble, c'est-à-dire $f(u) = 0$ si $u \in C$ et $f(u) = +\infty$ si $u \notin C$.

Le théorème de décomposition de tout à l'heure se généralise alors en invoquant la notion de *fonctions convexes duales* (ou *conjuguées*) introduite pour $H = \mathbf{R}$ par S. Mandelbrojt [10], puis pour $H = \mathbf{R}^n$ par W. Fenchel ([7], [8]), adaptée aux espaces de dimension infinie par A. Brøndsted [3] (comme sources plus lointaines et classiques, on pourrait également citer les *fonctions conjuguées de Young*, sur $[0, +\infty[$ ou même l'idée de *transformation de Legendre* en théorie élémentaire des équations aux dérivées partielles). Il nous a été nécessaire de reprendre cette notion sous une forme légèrement différente (fonction pouvant prendre la valeur $+\infty$, *cf.* [17]). Si $f \in \Gamma_0(H)$, la fonction

$$v \to g(v) = \sup_{u \in H} [(u \mid v) - f(u)]$$

est appelée *fonction duale* de $f$. Cette fonction appartient aussi à $\Gamma_0(H)$ et $f$ est, symétriquement, la fonction duale de $g$.

On peut dire aussi bien que $g$ est la plus petite fonction numérique telle que, pour tout $x$ et tout $y$ dans $H$, on ait

$$f(x) + g(y) \geqq (x \mid y).$$

Deux points $x$ et $y$ de $H$ sont déclarés *conjugués par rapport à $f$ et $g$* si

$$f(x) + g(y) = (x \mid y),$$

et la proposition clef du présent mémoire peut s'énoncer : *Tout $z \in H$ est égal à la somme de $x = \text{prox}_f z$ et de $y = \text{prox}_g z$; les points $x$ et $y$ sont conjugués par rapport à $f$ et $g$ et c'est là l'unique décomposition de $z$ en somme de deux éléments conjugués par rapport à $f$ et $g$.*

Ce résultat a déjà été publié dans [18]; dans [19] est paru un sommaire de certains développements ultérieurs, diffusés sous forme multigraphiée dans [20].

1.*c*. — Le présent article expose d'abord les éléments utiles pour la suite de la théorie des fonctions convexes duales. On reprend ensuite, avec des améliorations et compléments, tous les résultats du fascicule multigraphié [20], lequel est donc désormais périmé : existence de $\text{prox}_f z$, exemples élémentaires, théorème des fonctions duales, propriétés élémentaires des applications prox, fonction primitive d'une application prox, fonctions numériques « plus convexes » ou « moins convexes » que la fonction $z \to \frac{1}{2}\|z\|^2$ et caractérisations diverses des applications prox. La notion de « fonctions sur-duales » introduite dans la section 11 présente un intérêt technique pour les applications de la théorie à l'analyse fonctionnelle. On établit enfin un rapprochement entre la proposition 4.*a* et un théorème de G. J. MINTY sur les relations (ou applications multivoques) « monotoniques » dans un espace hilbertien.

Afin de ne pas trop allonger ce mémoire, nous avons réservé pour des publications ultérieures les applications essentielles de la théorie, où $H$ est réalisé comme espace fonctionnel (*voir* aussi [15], [27]).

1.*d*. — L'idée de problèmes variationnels « conjugués » qui domine la présente étude est très classique et a donné lieu à de nombreuses publications. Dans la plupart des cas, les éléments des deux problèmes sont, grossièrement parlant, mis en dualité, et il n'est pas essentiel que ces deux problèmes soient formulés dans un même espace fonctionnel. L'algorithme des fonctions convexes duales constitue justement un instrument très efficace pour l'exploitation systématique de cette idée dans le cas convexe (*cf.* C. BERGE, A. GHOUILA-HOURI [1], chap. 5 et R. T. ROCKAFELLAR [33] pour l'utilisation du procédé dans la théorie de la programmation convexe, en dimension finie). Il est beaucoup moins usuel de voir, comme ici, ces éléments directement comparés ou additionnés. Notons cependant quelque analogie entre notre paragraphe 5.*c* et les considérations mises par W. PRAGER et J. L. SYNGE [32], J. L. SYNGE [35] à la base de la « méthode de l'hypercercle » (pour le principe de cette méthode, *voir* COURANT et HILBERT [6], chap IV, § 11, et, pour un exposé technique détaillé, J. L. SYNGE [35]). Il s'agit chez ces auteurs de problèmes linéaires et les fondements géométriques consistent dans l'orthogonalité de variétés linéaires. Dans le même esprit, divers travaux ont été consacrés à la détermination de majorants et de minorants pour les valeurs extrémales des fonctionnelles quadratiques associées à des problèmes linéaires de la physique mathématique (pour un exposé de synthèse, *voir* J. B. DIAZ [7]) : notre paragraphe 9.*c* vise un but analogue dans le cas non linéaire.

## 2. Fonctions convexes duales.

2.*a*. — Nous présentons dans cette section les éléments utiles pour la suite de la théorie des *fonctions convexes duales*, telle qu'elle est exposée dans [17].

NOTATION. — *Notons $\Gamma(H)$ l'ensemble des fonctions numériques définies sur $H$ qui sont des enveloppes supérieures de familles de fonctions affines continues. Notons $\Gamma_0(H)$ l'ensemble des éléments de $\Gamma(H)$ autres que les deux constantes $-\infty$ (enveloppe supérieure d'une famille vide) et $+\infty$.*

A l'égard de cette notation, il est indifférent de considérer sur $H$ la topologie forte ou la topologie faible puisque la classe des formes linéaires continues (donc aussi la classe des fonctions affines continues) est la même pour ces deux topologies.

Il résulte des propriétés classiques des enveloppes supérieures que toute $f \in \Gamma_0(H)$ est une fonction à valeurs dans $]-\infty, +\infty]$, convexe et semi-continue inférieurement. Réciproquement, on montre (*cf.* [17], § 1, où l'on considère, plus généralement que l'hilbertien $H$, un espace vectoriel topologique localement convexe séparé quelconque) que *toute fonction à valeurs dans* $]-\infty, +\infty]$, *convexe et semi-continue inférieurement, autre que la constante* $+\infty$, *appartient à* $\Gamma_0(H)$. Ici aussi, *comme dans toute cette section*, il est indifférent de considérer sur $H$ la topologie forte ou la topologie faible, puisque la semi-continuité inférieure de la fonction *convexe* $f$ signifie que, pour tout $k \in \mathbf{R}$, l'ensemble *convexe* $\{x \in H : f(x) \leq k\}$ est fermé.

Exemple. — La *fonction indicatrice* $\psi_A$ d'une partie $A$ de $H$, c'est-à-dire

$$\psi_A(x) = \begin{cases} 0 & \text{si } x \in A, \\ +\infty & \text{si } x \notin A \end{cases}$$

appartient à $\Gamma_0(H)$ si et seulement si $A$ est convexe fermé non vide.

2.*b*. — Il résulte de la définition que toute $f \in \Gamma_0(H)$ est égale à l'enveloppe supérieure de la famille, certainement non vide, des fonctions affines continues qui la minorent.

Une fonction affine continue, soit $x \to (x \mid y) - \beta$, avec $y \in H$ et $\beta \in \mathbf{R}$ minore $f$ si et seulement si

$$\sup_{x \in H} [(x \mid y) - f(x)] \leq \beta, \tag{2.1}$$

ce qui conduit à s'intéresser à la fonction

$$y \to g(y) = \sup_{x \in H} [(x \mid y) - f(x)] = \sup_{f(x) < +\infty} [(x \mid y) - f(x)] \tag{2.2}$$

appelée *fonction duale* de $f$. Elle appartient aussi à $\Gamma_0(H)$, puisqu'elle est enveloppe supérieure d'une famille non vide de fonctions affines continues et qu'elle n'est pas partout égale à $+\infty$ [car (2.1) est satisfaite par au moins un $\beta \in \mathbf{R}$].

Pour construire $f$ comme enveloppe supérieure de ses minorantes affines continues; on peut évidemment retenir seulement celles de ces minorantes qui sont *maximales*, c'est-à-dire prendre $\beta = g(y)$, donc

$$f(x) = \sup_{g(y) < +\infty} [(x \mid y) - g(y)] = \sup_{y \in H} [(x \mid y) - g(y)],$$

ce qui signifie que *f est, symétriquement, la fonction duale de g*.

On vient de définir une bijection involutive de $\Gamma_0(H)$ sur lui-même. Cette involution se prolonge d'ailleurs à $\Gamma(H)$ car, selon la définition (2.2) les constantes $+\infty$ et $-\infty$ sont deux fonctions duales l'une de l'autre.

Exemple. — Prenons $f = \psi_C$, fonction indicatrice d'un ensemble $C$ convexe fermé non vide; la fonction duale

$$g(y) = \sup_{f(x) < +\infty} [(x \mid y) - f(x)] = \sup_{x \in C} (x \mid y)$$

est classiquement appelée *fonction d'appui* de l'ensemble $C$.

2.*c*. — La définition des fonctions duales implique

$$\forall x, \quad \forall y, \quad f(x) + g(y) \geqq (x \mid y). \tag{2.3}$$

Si deux points $x$ et $y$ sont tels qu'on ait l'égalité

$$f(x) + g(y) = (x \mid y), \tag{2.4}$$

on dit qu'ils sont *conjugués* par rapport au couple de fonctions duales $f$ et $g$. Cela signifie aussi que la fonction affine continue

$$u \to (u \mid y) - g(y) = (u - x \mid y) + f(x)$$

qui est une minorante de $f$, prend au point $x$ la même valeur que cette fonction (minorante de $f$ « exacte » au point $x$). Nous disons alors que $y$ est un *sous-gradient* de $f$ au point $x$.

Notation. — *L'ensemble des sous-gradients de $f$ au point $x$ est noté $\partial f(x)$*, de sorte que (2.4) peut s'écrire, de façon équivalente, $y \in \partial f(x)$ ou, symétriquement, $x \in \partial g(y)$.

L'ensemble $\partial f(x)$ est *convexe fermé* (éventuellement vide) puisqu'il s'écrit, d'après (2.3) et (2.4),

$$\begin{aligned} \partial f(x) &= \{ y \in H : f(x) + g(y) = (x \mid y) \} \\ &= \{ y \in H : g(y) - (x \mid y) \leqq - f(x) \} \end{aligned}$$

et que la fonction $y \to g(y) - (x \mid y)$ est convexe, semi-continue inférieurement.

2.*d*. — Si $\partial f(x) \neq \emptyset$ la fonction $f \in \Gamma_0(H)$ est dite *sous-différentiable* au point $x$; c'est évidemment le cas, en particulier, si cette fonction (convexe) est différentiable au sens de Gâteaux (différentiabilité « faible ») au point $x$ ou, *a fortiori*, au sens de Fréchet (différentiabilité « forte »); $\partial f(x)$ est alors réduit à un seul point. L'ensemble des points de $H$ où la fonction $f$ est sous-différentiable est appelé *assise* de $f$.

Pour une étude plus approfondie de la sous-différentiabilité dans les espaces vectoriels topologiques, on se reportera à [24], [25], [26].

L'utilité de la sous-différentiabilité tient, pour une part, à la remarque tautologique suivante : d'après la définition des fonctions duales, la borne inférieure de $f$ est $-g(0)$; l'ensemble des points où $f$ atteint cette borne est $\partial g(0)$; autrement dit encore $f$ présente un minimum au point $x$ si et seulement si $0 \in \partial f(x)$.

## 3. Applications « prox ».

3.*a*. Proposition. — *Soit* $f \in \Gamma_0(H)$; *quel que soit* $z \in H$, *la fonction numérique*

$$u \to \Phi(u) = \frac{1}{2} \| u - z \|^2 + f(u)$$

*possède un minimum strict.*

En effet, on remarquera d'abord que, de même que $f$, la fonction $u \to \frac{1}{2} \| u - z \|^2$ est semi-continue inférieurement sur $H$ faible puisque, pour tout $k \in \mathbf{R}$, la boule $\{ u \in H : \frac{1}{2} \| u - z \|^2 \leqq k \}$ est fermée dans $H$ fort ou faible; donc, la somme $\Phi$ est semi-continue inférieurement sur $H$ faible.

D'autre part, $f$ appartenant à $\Gamma_0(H)$ possède au moins une minorante affine continue, soit $u \to (a \mid u) - \beta$, de sorte que $\Phi(u)$ est minorée par l'expression

$$\frac{1}{2} \| u - z \|^2 + (a \mid u) - \beta = \frac{1}{2} \| u - (z - a) \|^2 + (a \mid z) - \frac{1}{2} \| a \|^2 - \beta.$$

Par hypothèse, il existe $b \in H$ tel que $f(b) \neq +\infty$. De la minoration précédente de $\Phi(u)$ il résulte qu'on peut trouver $\rho > 0$ tel que la condition $\| u - (z - a) \| > \rho$ entraîne $\Phi(u) > \Phi(b)$. Appelons $B$ une boule fermée de centre $z - a$, de rayon supérieur à $\rho$ et contenant $b$. On sait qu'une telle boule est compacte dans $H$ faible; donc $\Phi$, fonction numérique semi-continue inférieurement, présente sur cette boule un minimum en un point $x$. On a $\Phi(b) \geqq \Phi(x)$ et comme, hors de $B$, $\Phi$ surpasse $\Phi(b)$, ce minimum est aussi le minimum de $\Phi$ sur $H$.

Enfin, comme la fonction $u \to \frac{1}{2} \| u - z \|^2$ est *strictement convexe* et que $f$ est convexe, $\Phi$ est strictement convexe sur l'ensemble $\{ u \in H : \Phi(u) \neq +\infty \}$. Il en résulte que le minimum précédent est strict.

3.*b*. Notation. — *L'unique point* $x$ *où la fonction* $u \to \frac{1}{2} \| u - z \|^2 + f(u)$ *atteint son minimum sera noté* $x = \mathrm{prox}_f z$, « *point proximal* » *de* $z$ *relativement à la fonction* $f \in \Gamma_0(H)$.

Si $f$ est une fonction constante, $\mathrm{prox}_f$ est l'application identique.

3.*c*. Exemple. — Prenons pour $f$ la fonction affine continue

$$f(u) = (a \mid u) - \beta \qquad (\text{où } a \in H \text{ et } \beta \in \mathbf{R}).$$

Par le calcul déjà fait ci-dessus, on a

$$\Phi(u) = \frac{1}{2} \| u - (z - a) \|^2 + (a \mid z) - \frac{1}{2} \| a \|^2 - \beta$$

qui atteint son minimum pour $u = z - a$. Donc, ici

$$\text{prox}_f z = z - a.$$

Ainsi toute *translation* dans $H$ est une application prox.

3.*d*. Exemple. — Soit $C$ une partie *convexe fermée*, non vide de $H$, et prenons $f = \psi_C$ (*cf.* exemple 2.*a*). Dans ce cas $\Phi$ prend la valeur $+\infty$ hors de $C$; c'est donc quelque part sur $C$ que cette fonction présente son minimum; comme pour $u \in C$ on a $\Phi(u) = \frac{1}{2} \| u - z \|^2$, on voit que $\text{prox}_{\psi_C}$ est la *projection* de $z$ sur $C$ : point de $C$ le plus proche de $z$. En particulier, si $C$ est réduit à un unique point $c$, on a $\text{prox}_C z = c$ pour tout $z$.

3.*e*. Exemple. — Un cas usuel est la synthèse des deux précédents. En appelant encore $C$ un ensemble convexe fermé, on prend

$$f(u) = (a \mid u) - \beta + \psi_C(u),$$

c'est-à-dire que $f$ vaut $+\infty$ hors de $C$ et se réduit sur $C$ à une fonction affine continue. On trouve alors

$$\text{prox}_f z = \text{proj}_C (z - a).$$

3.*f*. Exemple. — Soit $k \geqq 0$ une constante et

$$f(u) = k \| u \|^2.$$

Un calcul élémentaire donne

$$\text{prox}_f z = \frac{1}{1 + 2k} z.$$

Donc toute *homothétie* (ici de centre O), de rapport $\lambda \in (0, 1)$ est une application prox.

Plus généralement, les résultats des paragraphes ultérieurs permettent facilement de montrer qu'une application *linéaire* de $H$ dans lui-même est une application prox si et seulement si elle est *autoadjointe, positive, de norme* $\leqq 1$.

### 4. Théorème des fonctions duales.

4.*a*. Proposition. — *Soient* $f \in \Gamma_0(H)$ *et* $g \in \Gamma_0(H)$ *deux fonctions duales l'une de l'autre; soient* $x$, $y$, $z$ *trois éléments de* $H$; *les propriétés* (I) *et* (II) *suivantes sont équivalentes :*

$$\text{(I)} \qquad z = x + y, \qquad f(x) + g(y) = (x \mid y);$$
$$\text{(II)} \qquad x = \operatorname{prox}_f z, \qquad y = \operatorname{prox}_g z.$$

*Démonstration.*

1° Soient $x$, $y$, $z$ possédant la propriété (I) [cela implique $f(x) \neq +\infty$]. Il résulte de la définition des fonctions duales que, pour tout $u \in H$ :

$$g(y) \geqq (u \mid y) - f(u),$$

donc

$$(x \mid y) - f(x) \geqq (u \mid y) - f(u).$$

Pour la fonction $\Phi$ du paragraphe 3.*a*, cela donne

$$\begin{aligned}\Phi(u) - \Phi(x) &= \frac{1}{2}\| u - x - y \|^2 + f(u) - \frac{1}{2}\| y \|^2 - f(x) \\ &\geqq \frac{1}{2}\| u - x - y \|^2 - \frac{1}{2}\| y \|^2 + (u \mid y) - (x \mid y)\end{aligned}$$

et ce dernier membre se réduit à $\frac{1}{2}\| u - x \|^2 \geqq 0$, nul seulement si $u = x$; donc $x = \operatorname{prox}_f z$. Raisonnement semblable pour $y$, lequel est donc égal à $\operatorname{prox}_g z$.

2° Réciproquement, soit $z \in H$, soit $x = \operatorname{prox}_f z$ [cela implique $f(x) \neq +\infty$] et soit $y' = z - x$. Pour tout $u \in H$ et pour tout $\lambda \in \left]0, 1\right[$, la convexité de $f$ donne

$$f[\lambda u + (1 - \lambda) x] \leqq \lambda f(u) + (1 - \lambda) f(x)$$

et, comme la fonction $\Phi(u)$ a son minimum au point $x$,

$$\frac{1}{2}\| \lambda u + (1 - \lambda) x - z \|^2 + f[\lambda u + (1 - \lambda) x] \geqq \frac{1}{2}\| x - z \|^2 + f(x).$$

En rapprochant ces deux inégalités, on obtient

$$\frac{1}{2}\| y' \|^2 - \frac{1}{2}\| \lambda(u - x) - y' \|^2 + f(x) \leqq \lambda f(u) + (1 - \lambda) f(x)$$

ou, après développement et réduction,

$$-\frac{1}{2}\lambda^2 \| u - x \|^2 + \lambda(u - x \mid y') \leqq \lambda f(u) - \lambda f(x).$$

Puisque $\lambda > 0$, cela équivaut à

$$(u \mid y') - f(u) \leqq (x \mid y') - f(x) + \frac{1}{2}\lambda \| u - x \|^2.$$

Comme $\lambda$ peut être pris arbitrairement voisin de zéro, il en résulte

$$(x \mid y') - f(x) = \max_{u \in H} [(u \mid y') - f(u)] = g(y').$$

Donc $x$, $y'$, $z$ possèdent la propriété (I); le 1° montre alors que $y' = \operatorname{prox}_g z$, ce qui achève la démonstration.

4.*b*. — Un cas particulier utile est celui où l'on prend pour $f$ la *fonction indicatrice* $\psi_P$ d'un *cône convexe fermé* $P$ (de sommet O). On voit immédiatement que la fonction duale

$$y \to g(y) = \sup_{x \in P} (x \mid y)$$

est la fonction indicatrice du cône convexe fermé

$$Q = \{ y \in H : (x \mid y) \leqq 0 \text{ pour tout } x \in P \}$$

dit *cône polaire* de $P$ (et la relation entre les deux cônes est symétrique).

D'après l'exemple 3.*d*, les applications $\operatorname{prox}_f$ et $\operatorname{prox}_g$ sont alors les *projections* $\operatorname{proj}_P$ et $\operatorname{proj}_Q$.

La relation des *points conjugués*

$$f(x) + g(y) = (x \mid y)$$

impliquant que $f(x)$ et $g(y)$ sont finis exige ici $f(x) = g(y) = 0$ : cette relation se réduit donc dans le cas présent à l'*orthogonalité* $(x \mid y) = 0$, avec $x \in P$ et $y \in Q$, d'où :

Corollaire. — *Si $P$ et $Q$ sont deux cônes mutuellement polaires dans $H$ et $x$, $y$, $z$ trois éléments de $H$, les propriétés* (I) *et* (II) *suivantes sont équivalentes :*

(I) $\quad z = x + y, \quad x \in P, \quad y \in Q, \quad (x \mid y) = 0;$

(II) $\quad x = \operatorname{proj}_P z, \quad y = \operatorname{proj}_Q z$

## 5. Contraction des distances.

5.*a*. — Remarquons que la relation « $x$ et $y$ conjugués par rapport au couple de fonctions duales $f$ et $g$ » est *monotonique* au sens de G. J. Minty [11], c'est-à-dire que, si $x'$, $y'$ est un autre couple de points conjugués par rapport à $f$ et $g$, on a

$$(x - x' \mid y - y') \geqq 0. \tag{5.1}$$

Cela s'obtient en écrivant la définition des points conjugués

$$\begin{aligned} f(x) + g(y) &= (x \mid y), \\ f(x') + g(y') &= (x' \mid y'), \end{aligned}$$

où tous les termes sont nécessairement finis, puis par l'inégalité (2.3)

$$\begin{aligned} (x \mid y') &\leqq f(x) + g(y'), \\ (x' \mid y) &\leqq f(x') + g(y). \end{aligned}$$

Ajouter membre à membre, simplifier, regrouper.

On va en tirer :

5.*b*. Proposition. — *Quels que soient z et z' dans H, on a*

$$\| \operatorname{prox}_f z - \operatorname{prox}_f z' \| \leqq \| z - z' \| \tag{5.2}$$

*de sorte que l'application* $\operatorname{prox}_f$ *est continue de H fort dans H fort.*

En effet, si l'on pose

$$\begin{aligned} x &= \operatorname{prox}_f z, & x' &= \operatorname{prox}_f z'; \\ y &= \operatorname{prox}_g z, & y' &= \operatorname{prox}_g z', \end{aligned}$$

on a, d'après la proposition 4.*a*,

$$\begin{aligned} \| z - z' \|^2 &= \| x + y - x' - y' \|^2 \\ &= \| x - x' \|^2 + \| y - y' \|^2 + 2(x - x' \mid y - y'). \end{aligned} \tag{5.3}$$

L'inégalité (5.1) entraîne donc

$$\| z - z' \|^2 \| \geqq \| x - x' \|^2.$$

On voit, en outre :

Remarque. — L'inégalité dans (5.2) a lieu si et seulement si $y - y' = 0$, c'est-à-dire si

$$\operatorname{prox}_g z = \operatorname{prox}_g z'.$$

5.*c*. — Cette contraction des distances est importante pour les techniques numériques. Le point $z$ étant donné, on cherche une construction approximative de $x = \operatorname{prox}_f z$. On construit (les considérations de la Section 11 pourront être utiles pour cela) un premier couple $x_0$, $y_0$ tel que $y_0 \in \partial f(x_0)$, puis la somme $z_0 = x_0 + y_0$; en vertu de la proposition 4.*a*, on a $x_0 = \operatorname{prox}_f z_0$, d'où

$$\| x - x_0 \| \leqq \| z - z_0 \|, \tag{5.4}$$

ce qui localise le point inconnu $x$ dans la boule de centre $x_0$, de rayon connu $\| z - z_0 \|$.

Une localisation plus raffinée peut être obtenue si l'on recherche simultanément $x = \mathrm{prox}_f z$ et $y = \mathrm{prox}_g z$. La relation (5.3) fournit en effet l'inégalité, plus serrée que (5.4),

$$\| x - x_0 \|^2 + \| y - y_0 \|^2 \leqq \| z - z_0 \|^2.$$

La connaissance de $y_0 \in \partial f(x_0)$ permet, d'autre part, d'améliorer l'approximation par une méthode de gradient.

## 6. Questions d'images.

6.*a*. PROPOSITION. — *L'image réciproque d'un point $a \in H$ par l'application* $\mathrm{prox}_f$ *est l'ensemble convexe fermé (éventuellement vide)*

$$\overset{-1}{\mathrm{prox}}_f(a) = a + \partial f(a).$$

En effet. pour que $\mathrm{prox}_f z = a$, il faut et il suffit, d'après la proposition 4.*a*, que $z = a + y$, où $y$ est conjugué de $a$ par rapport aux fonctions duales $f$ et $g$, ce qui se note aussi bien $y \in \partial f(a)$.

6.*b*. — L'image réciproque en question est non vide si et seulement si $\partial f(a)$ est non vide; donc :

COROLLAIRE. — *L'image* $\mathrm{prox}_f(H)$ *de l'espace entier est l'ensemble des points où la fonction f est sous-différentiable, ensemble appelé assise de f.*

Évidemment, si $f$ est sous-différentiable en un point $a$, on a $f(a) < +\infty$. Mais le contre-exemple suivant montre que $f(a) < +\infty$ n'implique pas nécessairement $\partial f(a) \neq \emptyset$

On prend $H = \mathbf{R}$ et

$$f(x) = \begin{cases} -2\sqrt{x} & \text{pour} \quad x \geqq 0, \\ +\infty & \text{pour} \quad x < 0, \end{cases}$$

ce qui donne comme fonction duale

$$g(y) = \begin{cases} +\infty & \text{pour} \quad y \geqq 0, \\ -\dfrac{1}{y} & \text{pour} \quad y < 0. \end{cases}$$

Bien que $f(0) \neq +\infty$, le point 0 ne possède pas de conjugué relativement à $f$ et $g$, puisqu'un tel conjugué, soit $y$, devrait vérifier $f(0) + g(y) = 0.y$, c'est-à-dire $g(y) = 0$. Or $g$ ne prend nulle part la valeur 0.

REMARQUE. — En rapprochant le corollaire ci-dessus de la proposition 3.*a* (existence de l'application prox), on voit que l'*assise de $f \in \Gamma_0(H)$ ne peut pas être vide.*

6.*c*. PROPOSITION. — *L'ensemble (éventuellement vide) des points fixes de l'application* $\operatorname{prox}_f$ *coïncide avec l'ensemble* $\partial g(0)$ *des points où la fonction* $f$ *atteint sa borne inférieure* (*cf.* remarque finale de la Section 2).

En effet, puisque

$$z = \operatorname{prox}_f z + \operatorname{prox}_g z,$$

on voit que $z$ est un point fixe de $\operatorname{prox}_f$ si et seulement si

$$\operatorname{prox}_g z = 0,$$

c'est-à-dire $z \in \partial g(0)$, d'après la proposition 6.*a*.

## 7. Fonction primitive d'une application prox.

7.*a*. — $\operatorname{prox}_f z$ étant, par définition, le point où la fonction

$$\Phi(u) = \frac{1}{2}\| u - z \|^2 + f(u)$$

atteint son minimum, il est naturel de s'intéresser à la valeur de ce minimum considérée comme fonction de $z$ :

DÉFINITION. — *La fonction, à valeurs dans* $\mathbf{R}$, *partout définie sur* $H$,

$$z \to \varphi(z) = \frac{1}{2}\| y \|^2 + f(x),$$

*où* $x = \operatorname{prox}_f z$ *et* $y = \operatorname{prox}_g z$, *est appelée fonction primitive de l'application* $\operatorname{prox}_g$.

La proposition 7.*d* légitimera cette dénomination.

7.*b*. PROPOSITION. — *La fonction* $\varphi$ *appartient à* $\Gamma_0(H)$ *et sa duale est la fonction*

$$y \to \theta(y) = g(y) + \frac{1}{2}\| y \|^2.$$

En effet, soit $\zeta$ un élément quelconque de $H$, soit $\xi = \operatorname{prox}_f \zeta$, soit $\eta = \operatorname{prox}_g \zeta$. La fonction affine continue.

$$z \to \mu_\zeta(z) = (\eta \mid z) - \frac{1}{2}\| \eta \|^2 - g(\eta)$$

est une minorante de $\varphi$, car

$$\varphi(z) - \mu_\zeta(z) = \frac{1}{2}\| y \|^2 + f(x) - (\eta \mid z) + \frac{1}{2}\| \eta \|^2 + g(\eta)$$

et, en vertu de l'inégalité

$$f(x) + g(\eta) - (\eta \mid x) \geqq 0,$$

cette différence est minorée par

$$\frac{1}{2}\| y \|^2 + \frac{1}{2}\| \eta \|^2 - (\eta \mid z - x) = \frac{1}{2}\| y - \eta \|^2 \geqq 0.$$

En outre, $\mu_\zeta$ est une minorante « exacte » de $\varphi$ : elle prend au point $\zeta$ la même valeur que $\varphi$ puisque, d'après la proposition 4.$a$, on a

$$\xi + \eta = \zeta,$$
$$f(\xi) + g(\eta) = (\xi \mid \eta),$$

ce qui donne

$$(7.1) \qquad \mu_\zeta(z) = (z - \zeta \mid \eta) + \varphi(\zeta).$$

Donc

$$(7.2) \qquad \varphi(z) = \sup_{\zeta \in H} \mu_\zeta(z),$$

cela montre que $\varphi$ appartient à $\Gamma_0(H)$, puisque enveloppe supérieure d'une famille de fonctions affines continues.

Pour chercher la duale de $\varphi$, on écrit (7.2) sous la forme

$$\begin{aligned} \varphi(z) &= \sup_{\eta \in \operatorname{prox}_g H} \left[ (\eta \mid z) - \frac{1}{2}\| \eta \|^2 - g(\eta) \right] \\ &\leqq \sup_{\eta \in H} [(\eta \mid z) - \theta(\eta)], \end{aligned}$$

c'est-à-dire que $\varphi$ est majorée par la duale de $\theta$; donc $\theta$ est majorée par la duale de $\varphi$. Il reste à prouver l'inégalité inverse :

$$\theta(y) \geqq \sup_{\zeta \in H} [(\zeta \mid y) - \varphi(\zeta)],$$

c'est-à-dire montrer que, pour tout $y$ et tout $\zeta$, la quantité

$$\theta(y) - (\zeta \mid y) + \varphi(\zeta) = g(y) + \frac{1}{2}\| y \|^2 - (\zeta \mid y) + f(\xi) + \frac{1}{2}\| \eta \|$$

est positive. De fait, en notant que

$$f(\xi) + g(y) \geqq (\xi \mid y),$$

on minore cette quantité par

$$\frac{1}{2}\| y \|^2 - (\zeta \mid y) + (\xi \mid y) + \frac{1}{2}\| \eta \|^2 = \frac{1}{2}\| \eta - y \|^2 \geqq 0$$

7.*c*. REMARQUE. — La proposition 7.*b* peut être rattachée à la théorie générale de l'*inf-convolution*. Avec la terminologie de [22] et [23], on reconnaît dans

$$\varphi(z) = \inf_{u \in H} \left[ \frac{1}{2} \| u - z \|^2 + f(u) \right]$$

l'inf-convolution de $f$ avec la fonction $\mathcal{Q} : z \to \frac{1}{2} \| z \|^2$. De même que $f$ et $\mathcal{Q}$, cette inf-convolution appartient à $\Gamma_0(H)$ parce que $\mathcal{Q}$ est *inf-compacte* (*cf.* [22], [23]) sur $H$ faible. Sa duale est la *somme* des duales de $f$ et de $\mathcal{Q}$, c'est-à-dire la fonction $g + \mathcal{Q} = \theta$, puisque — on le verra tout à l'heure — la duale de $\mathcal{Q}$ est $\mathcal{Q}$. C'est dans cet esprit qu'on peut adapter divers résultats du présent article à des espaces vectoriels topologiques plus généraux que l'hilbertien $H$.

7.*d*. PROPOSITION. — *En tout point $z$ de $H$ la fonction $\varphi$ est différentiable au sens de Fréchet; son gradient en ce point est $y = \mathrm{prox}_g z$.*

En effet, puisque la fonction $\mu_\zeta$, telle qu'elle est exprimée en (7.1), minore $\varphi$, on a

$$\varphi(z) - \varphi(\zeta) \geqq (z - \zeta \mid \eta), \tag{7.3}$$

avec $\eta = \mathrm{prox}_g \zeta$. De même, en échangeant $z$ et $\zeta$,

$$\varphi(\zeta) - \varphi(z) \geqq (\zeta - z \mid y), \tag{7.4}$$

avec $y = \mathrm{prox}_g z$. En rapprochant (7.3) et (7.4), on obtient

$$0 \leqq \varphi(\zeta) - \varphi(z) - (y \mid \zeta - z) \leqq (\eta - y \mid \zeta - z).$$

Comme $\eta$ et $y$ sont les images respectives de $\zeta$ et $z$ par l'application $\mathrm{prox}_g$, laquelle contracte les distances, le dernier membre admet la majoration

$$(\eta - y \mid \zeta - z) \leqq \| \eta - y \| \cdot \| \zeta - z \| \leqq \| \zeta - z \|^2.$$

Cela montre que la fonction $\varphi$ possède comme *différentielle de Fréchet* au point $z$ la forme linéaire

$$dz \to (y \mid dz).$$

Autrement dit, $y$ est le *gradient* de $\varphi$.

EXEMPLE. — Comme dans l'exemple 3.*d*, prenons $f = \psi_C$($C$, partie convexe fermée non vide de $H$). Alors

$$x = \mathrm{proj}_C z$$

de sorte que $\varphi(z)$ est la moitié du carré de la distance de $z$ à l'ensemble $C$. Il est élémentaire que cette fonction est différentiable au sens de Fréchet et que son gradient est $y = z - x$.

## 8. Détermination d'une fonction par son application prox.

8.*a*. — Comme on l'a déjà noté, l'application $\mathrm{prox}_g$ est inchangée si l'on ajoute à la fonction $g$ une constante finie; nous montrons la réciproque :

PROPOSITION. — *Si deux fonctions $g \in \Gamma_0(H)$ et $g' \in \Gamma_0(H)$ sont telles que* $\mathrm{prox}_g = \mathrm{prox}_{g'}$, *on a $g' = g + K$ ($K \in \mathbf{R}$ constante).*

En effet, d'après la proposition 7.*d*, les fonctions primitives $\varphi$ et $\varphi'$ des applications $\mathrm{prox}_g$ et $\mathrm{prox}_{g'}$ sont, dans ce cas, deux fonctions différentiables ayant même gradient; il existe donc $K \in \mathbf{R}$ telle que

$$\varphi' = \varphi - K,$$

d'où, en passant aux fonctions duales $\theta$ et $\theta'$,

$$\theta' = \theta + K.$$

Par la proposition 7.*b*, on en tire la conclusion cherchée.

8.*b*. De même, une fonction appartenant à $\Gamma_0(H)$ est déterminée, à une constante additive près, par son sous-gradient :

PROPOSITION. — *Si deux fonctions $g \in \Gamma_0(H)$ et $g' \in \Gamma_0(H)$ sont telles que pour tout $y \in H$, on ait $\partial g'(y) \subset \partial g(y)$, alors $g' = g + K$ ($K \in \mathbf{R}$ constante).*

En effet, la proposition 4.*a* montre que, pour tout $z \in H$, l'unique décomposition de la forme $z = x + y$, telle que $x \in \partial g(y)$ est obtenue pour $y = \mathrm{prox}_g z$. Donc $\mathrm{prox}_g = \mathrm{prox}_{g'}$, ce qui ramène à la proposition 8.*a*.

8.*c*. — Dans le même ordre d'idées, signalons que la présente théorie fournit des démonstrations simples pour les propositions suivantes :

1° *Toute fonction $f \in \Gamma_0(H)$ est déterminée par sa restriction à son assise : elle est égale à l'enveloppe supérieure des fonctions affines continues qui minorent cette restriction* [c'est-à-dire encore que $f$ est la plus grande fonction de $\Gamma_0(H)$ minorant ladite restriction].

2° *Toute fonction $f \in \Gamma_0(H)$ est égale à l'enveloppe supérieure de ses minorantes affines continues exactes* (c'est-à-dire, selon la terminologie déjà employée au paragraphe 2.*c*, minorantes prenant quelque part la même valeur que $f$).

En fait, il a été établi dans [24] que ces deux propriétés (d'ailleurs duales l'une de l'autre) sont vraies, plus généralement, lorsque, au lieu de l'hilbertien $H$, on considère un espace de Banach réflexif; A. BRØNDSTED et R. T. ROCKAFELLAR [4] transposant des idées de E. BISHOP et R. R. PHELPS [2] les ont ensuite étendues au cas d'un *espace de Banach quelconque.*

## 9. Fonctions plus convexes ou moins convexes que $\mathcal{Q}$.

9.*a*. Proposition. — *L'unique fonction qui soit égale à sa duale est*

$$\mathcal{Q}: \quad z \to \frac{1}{2}\| z \|^2.$$

En effet, l'identité

$$\operatorname{prox}_f z + \operatorname{prox}_g z = z$$

montre que, si les deux fonctions duales $f$ et $g$ sont égales, on a

$$\operatorname{prox}_f z = \operatorname{prox}_g z = \frac{1}{2} z.$$

La fonction primitive $\varphi$ de l'application $\operatorname{prox}_g$ est définie, à une constante additive près, par son gradient, ce qui donne ici

$$\varphi(z) = \frac{1}{4}\| z \|^2 + K \qquad (\text{avec } K \in \mathbf{R}).$$

On calcule alors la duale $\theta$ de $\varphi$ :

$$\begin{aligned}\theta(y) &= \sup_{z \in H}\left[(z \mid y) - \frac{1}{4}\| z \|^2 - K\right] \\ &= \sup_{z \in H}\left[2\left(\frac{1}{2}z \mid y\right) - \left\|\frac{1}{2}z\right\|^2 - K\right] \\ &= -K + \| y \|^2 + \sup_{z \in H}\left[-\left\| y - \frac{1}{2}z\right\|^2\right] \\ &= -K + \| y \|^2.\end{aligned}$$

La proposition 7.*b* implique donc

$$g(y) = \frac{1}{2}\| y \|^2 - K.$$

Par un calcul semblable à celui qu'on vient de faire, on constate que la duale de la fonction $y \to \frac{1}{2}\| y \|^2 - K$ est la fonction $z \to \frac{1}{2}\| z \|^2 + K$. Donc l'égalité $f = g$ a lieu si et seulement si $K = 0$, ce qui achève la démonstration.

9.*b*. Définition. — *Il s'agit ici, comme d'habitude, de fonctions à valeurs dans* $]-\infty, +\infty]$. *On dira qu'une fonction convexe* $\varphi_1$ *est plus convexe qu'une fonction convexe* $\varphi_2$ *(ou que* $\varphi_2$ *est moins convexe que* $\varphi_1$*) s'il existe une fonction convexe* $\gamma$ *telle que* $\varphi_1 = \varphi_2 + \gamma$.

PROPOSITION. — *Les trois propriétés suivantes sont équivalentes :*

(I) $\varphi \in \Gamma_0(H)$ *est moins convexe que* $\mathcal{Q}$;

(II) $\varphi \in \Gamma_0(H)$ *a pour duale une fonction plus convexe que* $\mathcal{Q}$;

(III) $\varphi$ *est la fonction primitive d'une application* prox.

Montrons d'abord que (I) $\Rightarrow$ (II). Par hypothèse, il existe une fonction convexe $\gamma$ telle que $\varphi + \gamma = \mathcal{Q}$. Comme $\mathcal{Q}$ est partout finie, il en est de même de $\varphi$ et $\gamma$ de sorte qu'on peut écrire

$$\gamma = \mathcal{Q} - \varphi;$$

$\mathcal{Q}$ est continue sur $H$ fort, $-\varphi$ est semi-continue supérieurement; donc $\gamma$ est semi-continue supérieurement sur $H$ fort. Puisque $\gamma$ est convexe, il en résulte (*cf.* BOURBAKI, *Espaces vectoriels topologiques*, chap. II, § 5, n° 2) que cette fonction est continue sur $H$ fort, donc qu'elle appartient à $\Gamma_0(H)$ Par suite, $\gamma$ est l'enveloppe supérieure de ses minorantes affines continues. Notons $(m_i)_{i \in I}$ cette famille de minorantes; on écrit

$$\varphi = \mathcal{Q} - \gamma = \mathcal{Q} - \sup_{i \in I} m_i = \inf_{i \in I} (\mathcal{Q} - m_i).$$

Les fonctions $\mathcal{Q} - m_i$ appartiennent à $\Gamma_0(H)$; il se trouve que leur enveloppe inférieure $\varphi$ appartient, par hypothèse, à $\Gamma_0(H)$; $\varphi$ est donc le plus grand élément de $\Gamma_0(H)$ minorant les $\mathcal{Q} - m_i$. Comme la dualité inverse l'ordre, la duale $\theta$ de $\varphi$ est le plus petit élément de $\Gamma_0(H)$ majorant les duales $\mathcal{Q} - m_i$, c'est-à-dire l'enveloppe supérieure de cette famille de duales [car l'enveloppe supérieure d'une famille de fonctions appartenant à $\Gamma_0(H)$ appartient aussi à $\Gamma_0(H)$]. Un calcul élémentaire fournit la duale de $\mathcal{Q} - m_i$ sous la forme $\mathcal{Q} + n_i$, où $n_i$ est une autre fonction affine continue. Par conséquent,

$$\theta = \sup_{i \in I} (\mathcal{Q} + n_i) = \mathcal{Q} + \sup_{i \in I} n_i,$$

ce qui montre bien que $\theta$ est la somme de $\mathcal{Q}$ et d'une fonction convexe.

Montrons ensuite que (II) $\Rightarrow$ (III). Soit $\theta \in \Gamma_0(H)$ une fonction plus convexe que $\mathcal{Q}$; alors $g = \theta - \mathcal{Q}$ est une fonction convexe à valeurs dans $]-\infty, +\infty]$, non partout égale à $+\infty$. Comme les deux termes $\theta$ et $-\mathcal{Q}$, la fonction $g$ est semi-continue inférieurement sur $H$ fort, donc elle appartient à $\Gamma_0(H)$. D'après la proposition 7. *b*, la fonction primitive $\varphi$ de l'application $\mathrm{prox}_g$ est la duale de $\theta = g + \mathcal{Q}$, ce qu'il fallait démontrer.

Montrons enfin que (III) $\Rightarrow$ (I). On suppose qu'il existe $g \in \Gamma_0(H)$ telle que $\varphi$ soit la fonction primitive de l'application $\mathrm{prox}_g$. On appelle $f$ la duale de $g$ et $\gamma$ la fonction primitive de l'application $\mathrm{prox}_f$. Soit $z$ quelconque dans $H$ et

$$x = \mathrm{prox}_f z, \qquad y = \mathrm{prox}_g z.$$

D'après la définition 7.*a*, on a

$$\varphi(z) = \frac{1}{2}\|y\|^2 + f(x),$$

$$\gamma(z) = \frac{1}{2}\|x\|^2 + g(y)$$

et, d'après la proposition 4.*a*,

$$z = x + y,$$
$$f(x) + g(y) = (x \mid y).$$

On en tire, par addition et réduction,

$$\varphi(z) + \gamma(z) = \mathcal{Q}(z). \tag{9.1}$$

Donc $\varphi$ est moins convexe que $\mathcal{Q}$.

9.*c*. Remarque. — L'identité (9.1) est importante pour les techniques numériques. Le point $z$ étant donné, il arrive qu'on s'intéresse plus à la valeur minimale $\varphi(z)$ de la fonction

$$u \to \frac{1}{2}\|z - u\|^2 + f(u)$$

qu'à la détermination précise du point $x = \operatorname{prox}_f z$, où ce minimum est atteint [l'estimation de la seule valeur numérique $\varphi(z)$ est d'ailleurs *a priori*, un problème beaucoup plus simple que la recherche de $x$, lequel est habituellement un élément d'un espace fonctionnel].

Le choix arbitraire d'un $u_0 \in H$ fournit d'abord, par le calcul de $\frac{1}{2}\|z - u_0\|^2 + f(u_0)$ un *majorant* de $\varphi(z)$. De même, si $v_0 \in H$, le calcul de $\frac{1}{2}\|z - v_0\|^2 + g(u_0)$ fournit un majorant de $\gamma(z)$, donc, par l'égalité (9.1) un *minorant* de $\varphi(z)$ qui se trouve ainsi encadré. Le calcul est d'ailleurs simplifié si l'on a pu choisir $u_0$ et $v_0$ conjugués (on trouvera au paragraphe 11.*b* des remarques techniquement utiles pour la construction d'un tel couple).

9.*d*. — De la caractérisation précédente des fonctions primitives d'application prox on peut tirer, par exemple :

Proposition. — *Soit* $(p_i)$, $i = 1, 2, \ldots, n$, *une famille finie d'applications* prox *et soit* $(\alpha_i)$, $i = 1, 2, \ldots, n$, *une famille de* $n$ *nombres réels positifs tels que* $\sum_{i=1}^{n} \alpha_i = k \leqq 1$. *Alors l'application*

$$z \to p(z) = \sum_{i=1}^{n} \alpha_i p_i(z)$$

*est une application* prox.

En effet, soit $\varphi_i$ la fonction primitive de l'application $p_i$. Alors $p(z)$ est le gradient au point $z$ de la fonction convexe

$$\varphi = \sum_{i=1}^{n} \alpha_i \varphi_i,$$

laquelle est moins convexe que $\mathcal{Q}$ puisqu'à chaque $\varphi_i$ s'associe $\gamma_i$ convexe telle que

$$\varphi_i + \gamma_i = \mathcal{Q},$$

ce qui donne

$$\varphi + \sum_{i=1}^{n} \alpha_i \gamma_i + (1-k)\mathcal{Q} = \mathcal{Q}.$$

## 10. Une caractérisation des applications prox.

10.*a*. — Soit $P$ une *application multivoque* : loi associant à tout $z \in H$ une partie $Pz$ (éventuellement vide) de $H$ (la donnée d'une telle application équivaut aussi bien à celle d'une relation binaire sur $H$, ou partie de l'espace produit $H \times H$). On dira que l'application $P$ *contracte les distances* si

$$x \in Pz, \qquad x' \in Pz' \quad \Rightarrow \quad \| x - x' \| \leqq \| z - z' \|.$$

Il en résulte évidemment que $z = z' \Rightarrow x = x'$ : l'ensemble $Pz$ contient donc au plus un élément; mais il peut être vide. L'application $P$ est « au plus univoque », ce qui revient à une application univoque non nécessairement partout définie sur $H$.

10.*b*. Proposition. — *Les trois propriétés formulées dans la proposition* 9.*b équivalent à la suivante :*

(IV) $\varphi \in \Gamma_0(H)$ *et l'application* (a priori *multivoque*) $z \to \partial\varphi(z)$ *contracte les distances.*

Il est clair en effet que (III) $\Rightarrow$ (IV) puisque, si $\varphi$ est la fonction primitive de l'application $\mathrm{prox}_g$, on a $\partial\varphi(z) = \{\mathrm{prox}_g z\}$ et qu'une application prox contracte les distances.

Montrons maintenant que (IV) $\Rightarrow$ (I).

Soit $\mathcal{F}$ la fonction primitive de l'application $\mathrm{prox}_\varphi$. Cette fonction est partout différentiable et son gradient au point $\zeta$ est

$$z = \mathrm{prox}_\varphi \zeta.$$

Il résulte de la proposition 4.*a* que si l'on pose $x = z - \zeta$, on a

$$x \in \partial\varphi(z).$$

De même, pour tout autre point $\zeta'$, avec $z' = \mathrm{prox}_\varphi \zeta'$, on a

$$z' - \zeta' = x' \in \partial\varphi(z').$$

Donc, si $\varphi$ possède la propriété (IV),

$$\| z' - \zeta' - z + \zeta \|^2 \leqq \| z' - z \|^2,$$

ce qui s'écrit encore

$$(10.1) \qquad (2z' - \zeta' - 2z + \zeta \mid \zeta' - \zeta) \geqq 0.$$

Or $2z - \zeta$ est le gradient au point $\zeta$ de la fonction $2\mathcal{F} - \mathcal{Q} = \mathcal{G}$. L'inégalité (10.1) montre que cette fonction est *convexe* (on en tire en effet que la restriction de la fonction $\mathcal{G}$ à toute droite de $H$ est une fonction d'une variable à dérivée croissante). Autrement dit, la fonction $2\mathcal{F}$ est *plus convexe que* $\mathcal{Q}$; par la proposition 9.*b*, il en résulte que sa duale est moins convexe que $\mathcal{Q}$. Or d'après la proposition 7.*b*, la duale de $\mathcal{F}$ est la fonction $\varphi + \mathcal{Q}$ : un calcul simple montre alors que la duale de $2\mathcal{F}$ est la fonction

$$w \to 2\varphi\left(\frac{1}{2}w\right) + 2\mathcal{Q}\left(\frac{1}{2}w\right).$$

Puisque cette duale est moins convexe que $\mathcal{Q}$, il existe une fonction convexe, notons-la $w \to 2\psi\left(\frac{1}{2}w\right)$, telle que

$$2\psi\left(\frac{1}{2}w\right) + 2\varphi\left(\frac{1}{2}w\right) + 2\mathcal{Q}\left(\frac{1}{2}w\right) = \mathcal{Q}(w).$$

En utilisant l'homogénéité de $\mathcal{Q}$ et en écrivant $u$ pour $\frac{1}{2}w$, cela donne

$$\varphi(u) + \psi(u) = \mathcal{Q}(u).$$

Donc $\varphi$ est moins convexe que $\mathcal{Q}$, ce qu'il fallait démontrer.

10.*c*. Corollaire. — *Une application p (univoque, partout définie) de H dans lui-même est une application* prox *si et seulement si elle contracte les distances et s'il existe une fonction (convexe)* $\varphi$ *telle que pour tout* $z \in H$ *on ait*

$$(10.2) \qquad p(z) \in \partial\varphi(z).$$

La condition est évidemment nécessaire d'après les propositions 5.*b* et 7.*d*. Réciproquement, si l'on a (10.2) pour tout $z \in H$ et si $p$ contracte les distances, le raisonnement du paragraphe 7.*d* prouve que $\varphi$ est partout différentiable et que $p$ est son gradient. Donc $\partial\varphi(z) = \{ p(z) \}$, ce qui ramène à la proposition précédente.

## 11. Fonctions sur-duales.

11.*a*. — Les résultats de cette section ont un intérêt technique pour les applications de la théorie à l'analyse fonctionnelle.

DÉFINITION 1. — *On dit que* F *et* G, *fonctions définies sur H, à valeurs dans* $[-\infty, +\infty]$ *sont sur-duales si, pour tout x et tout y dans H, on a*

$$F(x) + G(y) \geqq (x \mid y) \tag{11.1}$$

[*avec, si besoin est, la convention* $\infty - \infty = +\infty$, *de sorte que* (11.1) *équivaut à* $F(x) \geqq (x \mid y) - G(y)$].

DÉFINITION 2. — *Si F est une fonction, à valeurs dans* $[-\infty, +\infty]$, *partout définie sur H, on note* $F^*$ *la fonction*

$$F^*(y) = \sup_{x \in H} [(x \mid y) - F(x)]$$

*dite fonction polaire de F.*

Visiblement, $F^* \in \Gamma(H)$.

Si $F \in \Gamma(H)$, la fonction polaire $F^*$ est la *fonction duale de F* : l'usage, d'un vocable particulier pour ce cas insiste sur la réciprocité qui se manifeste alors entre les deux fonctions.

On voit que $F^*$ est *la plus petite fonction qui soit sur-duale de F* : deux fonctions sont sur-duales si et seulement si l'une quelconque d'entre elles majore la fonction polaire de l'autre.

Il résulte de la définition de $F^*$ qu'une fonction affine continue, soit $x \to (x \mid y) - \beta$, minore $F$ si et seulement si $\beta \geqq F^*(y)$. L'enveloppe supérieure de ces minorantes de $F$ peut donc s'écrire

$$x \to \sup_{y \in H} [(x \mid y) - F^*(y)] = F^{**}(x),$$

fonction polaire de $F^*$. Cette enveloppe supérieure est aussi *le plus grand élément* de $\Gamma(H)$ *minorant F* : nous l'appelons la Γ-régularisée de $F$.

Deux fonctions sur-duales $F$ et $G$ possèdent toujours au moins un *couple de minorantes duales*, c'est-à-dire un couple de deux fonctions $f \leqq F$ et $g \leqq G$ qui soient duales l'une de l'autre : tel est par exemple le couple des fonctions $F^{**}$ et $F^*$ ou, aussi bien, le couple des fonctions $G^*$ et $G^{**}$. D'ailleurs, le passage aux fonctions polaires inversant l'ordre,

$$f \leqq F \;\Rightarrow\; f^* = g \geqq F^* \;\Rightarrow\; f \leqq F^{**},$$
$$g \leqq G \;\Rightarrow\; g^* = f \geqq G^* \;\Rightarrow\; g \leqq G^{**},$$

d'où pour tout couple $f$, $g$ de minorantes duales des fonctions sur-duales *F et G*,

$$G^{\star} \leqq f \leqq F^{\star\star} \leqq F, \tag{11.2}$$

$$F^{\star} \leqq g \leqq G^{\star\star} \leqq G. \tag{11.3}$$

Les fonctions sur-duales $F$ et $G$ possèdent donc un seul couple de minorantes duales si et seulement si $F^{\star\star} = G^{\star}$ (ce qui équivaut à $F^{\star} = G^{\star\star}$).

11.*b*. Notation. — *Soit $F$ une fonction partout définie sur $H$, à valeurs dans $]-\infty, +\infty]$, autre que la constante $+\infty$. Pour tout $z \in H$, on notera $P_F z$ l'ensemble (éventuellement vide) des points où la fonction*

$$u \to \frac{1}{2} \| u - z \|^2 + F(u)$$

*atteint sa borne inférieure.*

Si $F$ est convexe, la fonction en question est strictement convexe, de sorte que $P_F z$ contient au plus un élément. Si $F \in \Gamma_0(H)$, on a $P_F z = \{ \operatorname{prox}_F z \}$, non vide quel que soit $z$.

Proposition. — *Soient $F$ et $G$ deux fonctions sur-duales; soient $f$ et $g$ un couple de minorantes duales. S'il existe $x$, $y$, $z$ tels que*

$$z = x + y,$$

$$F(x) + G(y) = (x \mid y), \tag{11.4}$$

*alors*

$$P_F z = \{ x \} = \{ \operatorname{prox}_f z \},$$

$$P_G z = \{ y \} = \{ \operatorname{prox}_g z \}.$$

*En outre, au point $x$,*

$$f(x) = F(x) = F^{\star\star}(x) = G^{\star}(x) \tag{11.5}$$

*et, au point $y$,*

$$g(y) = G(y) = G^{\star\star}(y) = F^{\star}(y). \tag{11.6}$$

*Démonstration.*

1° Notons d'abord que (11.4) implique que les valeurs $F(x)$ et $G(y)$ sont finies; donc ni $F$ ni $G$ n'est la constante $+\infty$; donc ni $G^{\star}$, ni $F^{\star}$ (qui respectivement les minorent) n'est la constante $+\infty$. Par suite, $G$ et $F$ possèdent des minorantes affines continues : ces deux fonctions prennent donc leurs valeurs dans $]-\infty, +\infty]$ et les conditions de la notation 11.*b* sont remplies.

2° Par hypothèse, $f \leqq F$ et $g \leqq G$, donc

$$f(x) + g(y) \leqq F(x) + G(y) = (x \mid y).$$

Puisque $f$ et $g$ sont duales, l'inégalité (2.3) exige alors

$$f(x) + g(y) = (x \mid y)$$

et la proposition 4.$a$ montre que

$$x = \mathrm{prox}_f z, \qquad y = \mathrm{prox}_g z.$$

3° D'autre part, (11.4) implique

$$F(x) \leqq \sup_{v \in H} [(x \mid v) - G(v)] = G^*(x).$$

Comme $F$ et $G$ sont sur-duales, $G^*$ minore $F$; cela exige qu'au point $x$, on ait

(11.7) $$F(x) = G^*(x).$$

Comme $G^*$ et $G^{**}$ constituent un couple de minorantes duales de $F$ et $G$, on a, d'après le 2°,

$$x = \mathrm{prox}_{G^*} z,$$

ce qui signifie que la fonction

$$u \to \frac{1}{2} \| u - z \|^2 + G^*(u)$$

présente un minimum strict au point $x$. Il en est donc de même de la fonction

$$x \to \frac{1}{2} \| u - z \|^2 + F(u)$$

qui la majore et prend la même valeur en ce point ; donc

$$P_F z = \{ x \}$$

et, symétriquement,

$$P_G z = \{ y \}.$$

4° Enfin, en rapprochant l'égalité (11.7) des inégalités (11.2), on trouve bien (11.5), puis, par symétrie, (11.6).

11.$c$. Proposition. — *Soit $F$, à valeurs dans $]-\infty, +\infty]$, convexe (de sorte que $P_F z$ est, soit vide, soit réduit à un élément). Si $P_F z = \{ x \}$, on a*

$$x = \mathrm{prox}_{F^{**}} z.$$

*Au point $x$ les deux fonctions $F$ et $F^{**}$ prennent la même valeur et admettent $z - x$ parmi leurs sous-gradients.*

En effet, le raisonnement utilisé au paragraphe 4.*a*, 2° (où l'on invoquait seulement la convexité de $f$, mais non sa semi-continuité inférieure) montre, en posant $y = z - x$, que

$$(x \mid y) - F(x) = F^{\star}(y), \tag{11.8}$$

d'où, par la proposition précédente,

$$x = \operatorname{prox}_{F^{\star\star}} z.$$

L'égalité $F(x) = F^{\star\star}(x)$ résulte de (11.5); alors (11.8) donne

$$F^{\star\star}(x) + F^{\star}(y) = (x \mid y),$$

c'est-à-dire $y \in \partial F^{\star\star}(x)$. Comme la fonction $F$ majore $F^{\star\star}$ et prend au point $x$ la même valeur, elle admet, elle aussi, $y$ comme sous-gradient en ce point.

## 12. Sur un théorème de G. J. Minty.

12.*a*. — Nous avons déjà rappelé au paragraphe 5.*a* que, selon G. J. Minty ([11], [12], [13], [14]) une relation binaire $\mathcal{R}$ sur $H$(ou, si l'on préfère, une partie de l'ensemble produit $H \times H$) est dite *monotonique* si l'on a l'implication

$$x \mathcal{R} y, \quad x' \mathcal{R} y' \quad \Rightarrow \quad (x - x' \mid y - y') \geqq 0.$$

(Antérieurement à G. J. Minty, cette hypothèse a été exploitée, pour une application univoque $x \to y$ de $H$ dans $H$, par E.H. Zarantonello [37]. Pour des développements plus récents, *voir* F. E. Browder [5], E. H. Zarantonello [39].)

La relation $\mathcal{R}$ est dite monotonique *maximale* si elle n'admet pas de prolongement strict qui soit encore monotonique (autrement dit, l'ensemble représentatif de $\mathcal{R}$ dans $H \times H$ n'est pas strictement contenu dans un autre ensemble monotonique). Un théorème de G. J. Minty peut alors se formuler comme suit :

*Si $\mathcal{R}$ est une relation monotonique maximale sur l'espace hilbertien $H$, tout $z \in H$ possède une unique décomposition $z = x + y$ telle que $x \mathcal{R} y$; les applications $z \to x$ et $z \to y$ ainsi définies sont continues.*

Le rapprochement s'impose avec notre proposition 4.*a*; le résultat suivant le précise :

12.*b*. Proposition. — *Si $f \in \Gamma_0(H)$, la relation $x \mathcal{R} y$ : $y \in \partial f(x)$ (c'est-à-dire : « $x$ et $y$ conjugués par rapport à et à sa fonction duale $g$ » est monotonique maximale.*

En effet, on a vu au paragraphe 5.$a$ que cette relation est monotonique. Pour prouver qu'elle est maximale, on doit montrer que si un couple $a \in H$, $b \in H$ assure l'implication

$$(12.1) \qquad y \in \partial f(x) \quad \Rightarrow \quad (x - a \mid y - b) \geqq 0,$$

alors $b \in \partial f(a)$.

Posons à cet effet

$$a' = \operatorname{prox}_f(a+b), \qquad b' = \operatorname{prox}_g(a+b).$$

La proposition 4.$a$ donne

$$(12.2) \qquad a' + b' = a + b\,;$$

$$(12.3) \qquad b' \in \partial f(a').$$

L'implication (12.1) exige donc

$$(a' - a \mid b' - b) \geqq 0,$$

c'est-à-dire, en tenant compte de (12.2),

$$\| a' - a \|^2 \leqq 0,$$

d'où

$$a = a', \qquad b = b',$$

ce qui, vu (12.3), achève la démonstration.

On voit que cette démonstration repose essentiellement sur la proposition 4.$a$; il ne semble donc pas que dans les pages qui précèdent, l'appel au théorème de Minty aurait permis de simplifier l'exposé.

## BIBLIOGRAPHIE.

[1] Berge (C.) et Ghouila-Houri (A.). — *Programmes, jeux et réseaux de transport.* — Paris, Dunod, 1962.

[2] Bishop (E.) and Phelps (R. R.). — The support functionals of a convex set, *Convexity*, p. 27-35. — Providence, American mathematical Society, 1963 (Proceedings of Symposia in pure Mathematics, 7).

[3] Brøndsted (Arne). — Conjugate convex functions in topological vector spaces, *Mat.-fys. Medd. Dansk. Vid. Selsk.*, t. 34, 1964, n° 2, 27 pages.

[4] Brøndsted (Arne) and Rockafellar (R. T.). — On the subdifferentiability of convex functions, *Proc. Amer. math. Soc.* (à paraître).

[5] Browder (Felix E.). — *Multi-valued monotone nonlinear mappings and duality mappings in Banach spaces* (à paraître).

[6] Courant (R.) and Hilbert (D.). — *Methods of mathematical physics,* Vol. 1. — New York, Interscience Publishers, 1953.

[7] Diaz (Joaquin B.). — Upper and lower bounds for quadratic integrals and at a point for solutions of linear boundary value problems, *Boundary problems in differential equations,* Proceedings of a symposium conducted by the Mathematics Research Center [1959, Madison], p. 47-83. — Madison, The University of Wisconsin Press, 1960.

[8] Fenchel (Werner). — *Convex cones, sets and functions. Lectures notes.* — Princeton, Princeton University Press, 1953.

[9] Fenchel (Werner). — On conjugate convex functions, *Canadian J. of Math.*, t. 1, 1949, p. 73-77.

[10] Mandelbrojt (Szolem). — Sur les fonctions convexes, *C. R. Acad. Sc.*, t. 209, 1939, p. 977-978.

[11] Minty (George J.). — Monotone (nonlinear) operators in Hilbert space, *Duke math. J.*, t. 29, 1962, p. 341-346.

[12] Minty (George J.). — On a « monotonicity » method for the solution of nonlinear equations in Banach spaces, *Proc. Nat. Acad. Sc. U. S. A.*, t. 50, 1963, p. 1038-1041.

[13] Minty (George J.). — On the monotonicity of the gradient of a convex function, *Pacific J. of Math.*, t. 14, 1964, p. 243-247.

[14] Minty (George J.). — On the solvability of nonlinear functional equations of « monotonic » type, *Pacific J. of Math.*, t. 14, 1964, p. 249-255.

[15] Moreau (Jean-Jacques). — Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires, *C. R. Acad. Sc.*, t. 255, 1962, p. 238-240.

[16] Moreau (Jean-Jacques). — *Sur la naissance de la cavitation dans un fluide incompressible*, Communication non publiée faite au Congrès international des Mathématiciens [1962, Stockholm].

[17] Moreau (Jean-Jacques). — Fonctions convexes en dualité, *Séminaire de Mathématiques de la Faculté des Sciences de Montpellier*, n° 1, 1962 (multigr.).

[18] Moreau (Jean-Jacques). — Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sc*, t. 255, 1962, p. 2897-2899.

[19] Moreau (Jean-Jacques). — Propriétés des applications « prox », *C. R. Acad. Sc.*, t. 256, 1963, p. 1069-1071.

[20] Moreau (Jean-Jacques). — Applications « prox », *Séminaire de Mathématiques de la Faculté des Sciences de Montpellier*, n° 2, 1963, 20 pages (multigr.).

[21] Moreau (Jean-Jacques). — Les liaisons unilatérales et le principe de Gauss, *C. R. Acad. Sc.*, t. 256, 1963, p. 871-874.

[22] Moreau (Jean-Jacques). — Inf-convolution, *Séminaire de Mathématiques de la Faculté des Sciences de Montpellier*, n° 3, 1963, 48 pages (multigr.).

[23] Moreau (Jean-Jacques). — Inf-convolution des fonctions numériques sur un espace vectoriel, *C. R. Acad. Sc.*, t. 256, 1963, p. 5047-5049.

[24] Moreau (Jean-Jacques). — Étude locale d'une fonctionnelle convexe, *Séminaire de Mathématiques de la Faculté des Sciences de Montpellier*, n° 5, 1963, 25 pages (multigr.).

[25] Moreau (Jean-Jacques). — Fonctionnelles sous-différentiables, *C. R. Acad. Sc.*, t. 257, 1963, p. 4117-4119.

[26] Moreau (Jean-Jacques). — Sur la fonction polaire d'une fonction semi-continue supérieurement, *C. R. Acad. Sc.*, t. 258, 1964, p. 1128-1130.

[27] Moreau (Jean-Jacques). — Sur la naissance de la cavitation dans une conduite, *C. R. Acad. Sc.*, t. 259, 1964, p. 3948-3950.

[28] Moreau (Jean-Jacques). — Semi-continuité du sous-gradient d'une fonctionnelle, *C. R. Acad. Sc.*, t. 260, 1965, p. 1067-1070.

[29] Prager (William). — Elastic solids of limited compressibility, *Actes du 9e Congrès international de Mécanique appliquée* [1957, Bruxelles], Vol. 5, p. 205-211. — Bruxelles, Université de Bruxelles, 1958.

[30] Prager (William). — Unilateral constraints in mechanics of continua, *Atti del Convegno Lagrangiano* [1963, Torino], p. 181-191. — Torino, Accademia delle Scienze, 1964 (Atti della Accademia della Scienze di Torino, t. 98, 1963-1964, Supplemento).

[31] PRAGER (William). — *On elastic, perfectly locking materials,* IBM Research Papers R. Z. 122, 1964.

[32] PRAGER (W.) and SYNGE (J. L.). — Approximations in elasticity based on the concept of function space, *Quarterly appl. Math.*, t. 5, 1947, p. 241-269.

[33] ROCKAFELLAR (R. T.). — *Convex functions and dual extremum problems,* Thèse Sc. math., Harvard University, 1963.

[34] ROCKAFELLAR (R. T.). — Duality theorems for convex functions, *Bull. Amer. math. Soc.*, t. 70, 1964, p. 189-192.

[35] SYNGE (John L.). — The method of the hypercircle in function-space for boundary value problems, *Proc. Royal Sc. London*, Series A, t. 191, 1947, p. 447-467.

[36] SYNGE (John L.). — *The hypercircle in mathematical physics.* — Cambridge, Cambridge University Press, 1957.

[37] ZARANTONELLO (Eduardo H.). — *Solving functional equations by contractive averaging.* — Madison, University of Wisconsin, 1960 (Mathematics Research Center U. S. Army, Technical Reports n° 160).

[38] ZARANTONELLO (Eduardo H.). — *The closure of the numerical range contains the spectrum.* — Lawrence, University of Kansas, 1964 (University of Kansas, Department of Mathematics, Technical Reports n° 7).

[39] ZARANTONELLO (Eduardo H.). — The closure of the numerical range contains the spectrum, *Bull. Amer. math. Soc.*, t. 70, 1964, p. 781-787.

(Manuscrit reçu le 22 janvier 1965.)

Jean-Jacques MOREAU,
Professeur à la Faculté des Sciences,
Chemin des Brusses,
34-Montpellier (Hérault).