HAL open science

# Sous-différentiabilité

Jean Jacques Moreau

▶ To cite this version:

Jean Jacques Moreau. Sous-différentiabilité. Colloquium on Convexity, 1967, Copenhague, Danemark. pp.185-201. hal-02161970

HAL Id: hal-02161970
https://hal.science/hal-02161970v1

Submitted on 21 Jun 2019

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

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

# SOUS-DIFFERENTIABILITE

J. J. MOREAU

## 1. Introduction.

**1.a.** La notion de *sous-différentiabilité* des fonctions numériques définies sur un espace vectoriel topologique est étroitement reliée à la théorie des *fonctions convexes conjuguées*, inaugurée par W. Fenchel. Durant les trois dernières années cette théorie a été assez largement développée par A. Brøndsted, U. Dieter, R. T. Rockafellar et l'auteur. Les résultats ont été publiés de manière dispersée, avec des améliorations successives, souvent sous forme d'exposés de séminaires multigraphiés ou de « preprints » d'articles non encore imprimés.

Nous voulons ici résumer les principaux résultats acquis en indiquant où trouver les démonstrations.

Les applications de cette théorie concernent essentiellement la recherche d'extremums mais sont de nature très diverse: elles vont de la programmation mathématique (cf. Rockafellar [46], [47], [52], [53], Dieter [16]) jusqu'aux problèmes d'analyse fonctionnelle posés par les systèmes mécaniques à liaisons unilatérales (cf. Moreau [29], [39], [42], [43], [44]). D'autre part, les applications sous-différentielles, qui font l'objet du paragraphe 9, fournissent l'exemple le plus simple d'applications (multivoques) d'un espace vectoriel topologique dans son dual qui sont *monotones* au sens de G. J. Minty. Cette dernière notion prend actuellement une grande importance en analyse fonctionnelle, avec notamment des travaux de F. E. Browder, J. Leray, J. L. Lions, G. J. Minty, G. Stampacchia, M. I. Visik. Le cas particulier des applications sous-différentielles peut servir de guide dans l'étude des applications monotones plus générales (cf. Aggeri, Lescarret [2], Lescarret [21]).

**1.b.** Nous considérons dans toute la suite un couple d'espaces vectoriels $X$ et $Y$, sur le corps $\boldsymbol{R}$ des réels, *mis en dualité* par la forme bilinéaire $\langle .,. \rangle$ (cf. Bourbaki [11]). L'exemple le plus usuel de cette situation est obtenu en prenant pour $X$ un espace vectoriel, muni a priori d'une topologie localement convexe séparée, et pour $Y$ son dual topologique, muni de la topologie faible. On sait que chacun des deux espaces — soit $X$ pour fixer les idées — peut être muni de topologies localement

convexes séparées qui sont *compatibles avec la dualité* en ce sens que, pour ces topologies, la famille des formes linéaires continues est identique à la famille des fonctions de la forme $\langle .,y\rangle$ avec $y \in Y$. La *topologie de Mackey* $\tau(X,Y)$ est la plus fine de ces topologies, la *topologie faible* $\sigma(X,Y)$ est la moins fine; elles se confondent si les espaces sont de dimension finie.

**1.c.** Rappelons que la famille des *ensembles convexes fermés* dans $X$ (resp. dans $Y$) est la même pour toutes les topologies compatibles avec la dualité; il en est ainsi également de la famille des *ensembles bornés*. C'est en ce sens qu'on parlera dans la suite d'ensembles convexes fermés, ou d'ensembles bornés, sans spécifier une topologie.

De la notion d'ensemble borné dans l'un des deux espaces $X$ ou $Y$ résulte la définition de la *topologie forte* sur l'autre: $X$ étant interprété comme ensemble de formes linéaires sur $Y$, la topologie forte sur $X$ est, par définition, la topologie (localement convexe, séparée) de la convergence uniforme dans les parties bornées de $Y$; nous la noterons $\beta(X,Y)$. Elle est visiblement plus fine que $\tau(X,Y)$, laquelle est la topologie de la convergence uniforme dans les parties convexes faiblement compactes de $Y$. Par conséquent $\beta(X,Y)$ est compatible avec la dualité seulement si elle coincide avec $\tau(X,Y)$. Le cas *réflexif* est celui où on a simultanément $\beta(X,Y)=\tau(X,Y)$ et $\beta(Y,X)=\tau(Y,X)$.

## 2. Sous-gradients d'une fonction en un point.

**2.a.** Soit $h$ une fonction définie, pour fixer les idées, sur l'espace $X$, à valeurs dans $\bar{\boldsymbol{R}}=[-\infty, +\infty]$. Comme il sera essentiellement question de *minorations* de $h$, le cas d'une fonction qui serait définie seulement sur une partie de $X$ se raménera immédiatement à celui-là en prolongeant la fonction avec la valeur $+\infty$ hors de son domaine initial.

DÉFINITIONS. *On dit que $y \in Y$ est un sous-gradient de $h \in \bar{\boldsymbol{R}}^X$ au point $x_0 \in X$ si la valeur $h(x_0)$ est finie et si la fonction affine continue*

$$x \to \langle x-x_0, y\rangle + h(x_0)$$

*minore $h$ partout sur $X$.*

Cette minorante prenant au point $x_0$ la même valeur que $h$, nous disons qu'elle est *exacte au point $x_0$*.

*On appelle sous-différentiel de $h$ au point $x_0$ l'ensemble des sous-gradients de $h$ en ce point.*

C'est une partie, éventuellement vide, de $Y$ qu'on notera $\partial h(x_0)$.

*Si cet ensemble n'est pas vide, la fonction $h$ est dite sous-différentiable au point $x_0$.*

Ces notions tirent une partie de leur intérêt de la remarque tautologique suivante:

*La fonction $h$ possède un minimum, atteint au point $x_0$, si et seulement si le sous-différentiel $\partial h(x_0)$ contient le zéro de $Y$.*

**2.b.** Exemple. *Soit $A$ une partie de $X$; la fonction numérique $\psi_A$ définie par*

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

*sera appelée fonction indicatrice de l'ensemble $A$.*

Evidemment, pour tout $x_0 \in A$ l'ensemble $\partial\psi_A(x_0)$ contient au moins le zéro de $Y$; ce sous-différentiel contiendra un élément non nul $y_0$ si et seulement si l'ensemble $A$ possède au point $x_0$ un *hyperplan fermé d'appui* et l'équation de cet hyperplan s'écrira

$$\langle x-x_0, y_0\rangle = 0 .$$

Par là, l'étude des hyperplans d'appui de $A$ est ramené à la sous-différentiation de la fonction $\psi_A$.

**2.c.** Inversement, d'ailleurs, on peut associer à toute fonction numérique $h \in \bar{\boldsymbol{R}}^X$ son *épigraphe* dans l'espace produit $X \times \boldsymbol{R}$ c'est à dire l'ensemble

$$H = \{(x,r) \in X \times \boldsymbol{R}:\ h(x) \leqq r\} .$$

Et $y_0 \in Y$ est sous-gradient de $h$ au point $x_0$ si et seulement si, dans l'espace vectoriel topologique $X \times \boldsymbol{R}$, l'hyperplan fermé « non vertical »

$$r = \langle x-x_0, y_0\rangle + h(x_0)$$

est hyperplan d'appui de $H$ au point $(x_0, h(x_0))$.

Notons, dans cet ordre d'idées, que $h \in \bar{\boldsymbol{R}}^X$ est convexe (resp. semi-continue inférieurement) si et seulement si son épigraphe est un ensemble convexe (resp. fermé) dans $X \times \boldsymbol{R}$ (sur la notion de convexité pour des fonctions prenant la valeur $-\infty$, voir Moreau [33], [34]).

## 3. Fonction polaire d'une fonction numerique.

**3.a.** Soit la fonction affine continue sur $X$

$$x \to \langle x,y\rangle - r \tag{3.1}$$

avec $r \in \boldsymbol{R}$ et $y \in Y$; on dira que $y$ est la *pente* de cette fonction affine. La fonction ainsi définie est une minorante de $h \in \bar{\boldsymbol{R}}^X$ si et seulement si

$$r \geqq \sup\{\langle x,y\rangle - h(x) : x \in X\} \tag{3.2}$$

ce qui conduit à s'intéresser à la fonction $h^*$ définie sur $Y$ par le second membre de (3.2), c'est à dire aussi bien

$$h^*(y) = \sup\{\langle x,y\rangle - h(x) : h(x) < +\infty\} \tag{3.3}$$

que nous appelons *fonction polaire* de $h$.

D'après (3.3), $h^*$ est l'enveloppe supérieure d'une famille de fonctions affines continues sur $Y$: elle est donc *convexe* et *semi-continue inférieurement* sur l'espace $Y$ (pour n'importe laquelle des topologies compatibles avec la dualité). Elle prend ses valeurs dans $]-\infty, +\infty]$, à moins que $h$ ne soit la fonction constante $\omega = +\infty$, auquel cas $h^*$ est la fonction constante $-\omega = -\infty$.

**3.b.** Si $y$ est tel que $h^*(y) \in \boldsymbol{R}$ et si on prend $r = h^*(y)$, la minorante affine (3.1) est *maximale*. Et cette minorante maximale de $h$ est *exacte* (§ 2.a) en un point $x_0 \in X$ si et seulement si le sup dans (3.2) est un maximum, atteint pour $x = x_0$, c'est à dire si

$$h(x_0) + h^*(y) - \langle x_0, y\rangle = 0\,. \tag{3.4}$$

On tire de là une expression du sous-différentiel de $h$ au point $x_0$:

$$\begin{aligned} \partial h(x_0) &= \{y \in Y : h(x_0) + h^*(y) - \langle x_0, y\rangle = 0\} \\ &= \{y \in Y : h(x_0) + h^*(y) - \langle x_0, y\rangle \leqq 0\} \end{aligned} \tag{3.5}$$

qui montre que cet ensemble (éventuellement vide) est *convexe et fermé* dans $Y$.

**3.c.** Exemple. Si on prend pour $h$ la fonction indicatrice (§ 2.b) $\psi_A$ d'un ensemble $A \subset X$, la fonction polaire s'écrit, d'après (3.3),

$$\psi_A{}^*(y) = \sup\{\langle x,y\rangle : x \in A\}$$

expression bien connue sous le nom (assez impropre) de *fonction d'appui* de l'ensemble $A$.

**3.d.** Parmi les règles de calcul des fonctions polaires, la plus simple concerne la *translation*. Si $h \in \bar{\boldsymbol{R}}^X$ et $a \in X$, on note classiquement $T_a h$ la fonction *translatée*

$$x \to h(x-a)$$

on trouve alors

$$(T_a h)^* = h^* + \langle a, . \rangle . \tag{3.6}$$

De même si on ajoute à $h$ une forme linéaire continue, la fonction polaire subit une translation dans $Y$ et, plus généralement, si $b \in Y$ et $s \in \boldsymbol{R}$, on a

$$(h + \langle ., b \rangle - s)^* = s + T_b(h^*) . \tag{3.7}$$

Ces propriétés sont techniquement utiles pour ramener à l'origine un point en lequel on procède à une étude locale.

**3.e.** On appelle *inf-convolution* la loi de composition, commutative et associative, faisant correspondre à tout couple de fonctions $h_1$, $h_2$ définies sur $X$, à valeurs dans $\bar{\boldsymbol{R}}$, la fonction

$$(h_1 \nabla h_2)(x) = \inf\{h_1(u_1) + h_2(u_2) : u_1 + u_2 = x\} \tag{3.8}$$

(la somme $h_1(u_1) + h_2(u_2)$ est ici entendue avec la convention $(+\infty) + (-\infty) = (+\infty)$; cf. Moreau [31], [32], [33], [34]).

Exemple 1. Si on prend pour $h_1$ et $h_2$ les *fonctions indicatrices* de deux parties $A$ et $B$ de $X$, la fonction $\psi_A \nabla \psi_B$ est la fonction indicatrice de l'ensemble *somme* $A + B$.

Exemple 2. Si on note $\|.\|$ une norme sur $X$, l'expression $(\psi_A \nabla \|.\|)(x)$ représente la *distance* de $x$ à l'ensemble $A$.

L'inf-convolution nous intéresse ici par la propriété suivante des fonctions polaires:

$$(h_1 \nabla h_2)^* = h_1^* + h_2^* . \tag{3.9}$$

## 4. Cônes fonctionnels $\Gamma$ et fonctions duales.

**4.a.** Définition. *Si $E$ est un espace vectoriel topologique, nous notons $\Gamma(E)$ l'ensemble des $f \in \bar{\boldsymbol{R}}^E$ qui sont des enveloppes supérieures de fonctions affines continues. Nous notons $\Gamma_0(E)$ l'ensemble des éléments de $\Gamma(E)$ autres que les deux constantes $\omega = +\infty$ et $-\omega = -\infty$ (enveloppe supérieure d'une famille vide).*

L'enveloppe supérieure de toute famille de fonctions de $\Gamma(E)$ appartient aussi à $\Gamma(E)$. Il en résulte que $\Gamma(E)$, ordonné par l'ordre classique de $\bar{\boldsymbol{R}}^E$, est *complètement réticulé*: toute partie $G$ de $\Gamma(E)$ possède un sup, égal à son sup dans $\bar{\boldsymbol{R}}^E$, et un inf, noté $\Gamma \cdot \inf G$ (ce dernier est généralement plus petit que l'inf de $G$ dans $\bar{\boldsymbol{R}}^E$).

Pour toute fonction numérique $h \in \bar{\boldsymbol{R}}^E$, *l'enveloppe supérieure des éléments de $\Gamma(E)$ minorant $h$* (c'est aussi *le plus grand* élément de $\Gamma(E)$ minorant $h$) est appelée *$\Gamma$-régularisée* de $h$.

L'ensemble $\Gamma_0(E)\cup\{\omega\}$ est stable pour les opérations d'addition et de multiplication par toute constante $\lambda \in ]0, +\infty[$, ce qui munit cet ensemble d'une structure de *cône convexe*. Cette structure est étendue à $\Gamma(E)$ lui-même si on pose la convention $(+\infty)+(-\infty)=-\infty$ (pour l'usage systématique de cette *addition inférieure* voir Moreau [33], [34]).

**4.b.** Evidemment tout élément de $\Gamma(E)$ est une fonction *convexe* et *semi-continue inférieurement* sur $E$; réciproquement en utilisant les théorèmes classiques de séparation (c'est à dire, si on veut, le théorème de Hahn–Banach), on trouve:

*Si la topologie de $E$ est localement convexe séparée, toute fonction, à valeurs dans $]-\infty, +\infty]$, qui est convexe et semi-continue inférieurement appartient à $\Gamma(E)$.*

Il résulte de là que, $E$ étant supposé localement convexe séparé, si on considère une fonction $h$ *convexe*, à valeurs dans $]-\infty, +\infty]$ et sa *régularisée semi-continue inférieurement* $\hat{h}$, définie par

$$\hat{h}(x) = \liminf_{u\to x} h(u)$$

et si cette fonction $\hat{h}$ ne prend nulle part la valeur $-\infty$, elle est égale à la *$\Gamma$-régularisée* de $h$. Le cas usuel permettant d'affirmer que $\hat{h}$ ne prend nulle part la valeur $-\infty$ est celui où on possède une minorante de $h$ qui soit une fonction semi-continue inférieurement, à valeurs dans $]-\infty, +\infty]$ (par exemple une minorante appartenant à $\Gamma_0(E)$ ou, simplement, une minorante affine continue).

**4.c.** Si on prend pour $E$ l'un ou l'autre des espaces vectoriels en dualité $X$ et $Y$, les ensembles $\Gamma(X)$ et $\Gamma(Y)$ sont les mêmes pour toutes les topologies compatibles avec la dualité.

Il résulte de (3.3) que, pour toute $h \in \overline{\boldsymbol{R}}^X$, la fonction polaire $h^*$ appartient à $\Gamma(Y)$. La polaire $h^{**}$ de cette fonction polaire appartient à $\Gamma(X)$: c'est précisément la *$\Gamma$-régularisée* de $h$.

*En particulier lorsqu'on suppose a priori $f \in \Gamma(X)$ et $g \in \Gamma(Y)$ on a l'équivalence:*

$$f^* = g \iff f = g^* . \tag{4.1}$$

*Dans ce cas nous disons que $f$ et $g$ constituent une paire de fonctions duales.*

Autrement dit, le couple $f,g$ constitue alors un *élément minimal* dans l'ensemble des couples $f \in \overline{\boldsymbol{R}}^X$, $g \in \overline{\boldsymbol{R}}^Y$ tels que pour tout $x \in X$ et tout $y \in Y$ on ait

$$f(x)+g(y)-\langle x,y\rangle \geqq 0, \tag{4.2}$$

ensemble ordonné par l'ordre naturel de $\bar{\boldsymbol{R}}^X \times \bar{\boldsymbol{R}}^Y$ (l'addition est ici une *addition supérieure*: $(+\infty)+(-\infty)=(+\infty)$).

La dualité définit par là une bijection (renversant l'ordre) entre $\Gamma(X)$ et $\Gamma(Y)$. Comme d'ailleurs la fonction $\omega_X$ (constante $+\infty$ sur $X$) a pour duale la fonction $-\omega_Y$ (constante $-\infty$ sur $Y$) et inversement, la correspondance est aussi bijective entre $\Gamma_0(X)$ et $\Gamma_0(Y)$ (ensembles des « proper closed convex functions »).

Rappelons que, pour le cas $X=Y=\boldsymbol{R}^n$, avec des fonctions convexes à valeurs finies, la présente idée fut inaugurée dans Fenchel [17], [18]. Comme source plus ancienne, on peut citer les *fonctions conjuguées de Young*, définies sur $[0, +\infty[$ ou les couples de fonctions définies de manière analoge sur $\boldsymbol{R}$ dans Birnbaum–Orlicz [9] ou Mandelbrojt [22]. (Noter que des fonctions continues définies sur le coin positif de $\boldsymbol{R}^n$ et généralisant à certains égards les conjuguées de Young ont été considérées dans Bellman–Karush [3], [4], [5], [6], [7].)

L'idée peut également être rattachée à la *transformation de Legendre*, usuelle dans la théorie élémentaire des équations aux dérivées partielles: voir à cet égard Rockafellar [54].

**4.d.** L'addition dans $\Gamma(Y)$ se transmue par dualité en une loi de composition, commutative et associative, dans $\Gamma(X)$ que nous appelons *$\Gamma$-convolution*. On tire immédiatement de (3.9) que

$$(f_1{}^*+f_2{}^*)^* = (f_1 \nabla f_2)^{**} .$$

Autrement dit, la *$\Gamma$-convolution* est la *$\Gamma$-régularisée* de l'*inf–convolution*. Les cas où l'on a $f_1 \nabla f_2 \in \Gamma(X)$, de sorte que la $\Gamma$-convolution coïncide avec l'inf–convolution, sont d'une grande importance pratique (notamment dans l'étude de la sous-différentiabilité) ainsi que les cas où, dans la définition (3.8) de $h_1 \nabla h_2$, l'inf est un *minimum* (cf. Brøndsted [12], Moreau [31], [32], Rockafellar [46], [49].

**4.e.** *Deux éléments $x \in X$ et $y \in Y$ sont dits conjugués par rapport au couple de fonctions duales $f,g$ si on a l'égalité dans* (4.2):

$$f(x)+g(y) = \langle x,y \rangle . \tag{4.3}$$

D'après (3.5) cela équivaut à

$$y \in \partial f(x)$$

ou, tout aussi bien, à

$$x \in \partial g(y) .$$

**4.f.** Pour toute fonction numérique $h \in \bar{\boldsymbol{R}}^X$ on a les implications

$$\partial h(x_0) \neq \emptyset \Rightarrow h(x_0) = h^{**}(x_0)\,, \tag{4.4}$$

$$h(x_0) = h^{**}(x_0) \Rightarrow \partial h(x_0) = \partial h^{**}(x_0)\,, \tag{4.5}$$

$$y_0 \in \partial h(x_0) \Rightarrow x_0 \in \partial h^*(y_0)\,. \tag{4.6}$$

Par (4.4) et (4.5), l'étude du sous-différentiel d'une fonction $h$ en un point est, dans une large mesure, ramenée à celle du sous-différentiel de sa $\Gamma$-régularisée $h^{**}$. Cela justifie l'intérêt porté dans la suite aux fonctions appartenant à $\Gamma$.

## 5. Différentiabilité classique et sous-différentiabilité.

**5.a.** On sait qu'une fonction $f \in \bar{\boldsymbol{R}}^X$, finie au point $x_0$, est dite *différentiable au sens de Gâteaux* (ou faiblement différentiable) s'il existe $b \in Y$ (évidemment unique) tel que, pour tout $u \in X$ la fonction

$$t \to f(x_0 + tu) \tag{5.1}$$

de la variable réelle $t$ admette pour $t = 0$ une dérivée égale à

$$\frac{d}{dt} f(x_0 + tu)\,|_{t=0} = \langle u, b \rangle\,.$$

C'est ce qui a lieu, a fortiori, lorsque, $X$ étant un espace normé, la fonction est différentiable *au sens de Fréchet*.

On dit usuellement que $b$ est le *gradient* de $f$ au point $x_0$.

Si, en outre, la fonction $f$ est *convexe* on voit immédiatement qu'elle est, dans ce cas, sous-différentiable, et que son sous-différentiel $\partial f(x_0)$ se réduit au simple élément $b$.

**5.b.** Réciproquement, la théorie de la sous-différentiabilité permet, dans certains cas, d'établir la différentiabilité au sens Gâteaux:

Supposons que la fonction *convexe* $f \in \bar{\boldsymbol{R}}^X$ soit *finie et continue* au point $x_0$ (pour une topologie localement convexe compatible avec la dualité entre $X$ et $Y$, donc, a fortiori, pour la topologie de Mackey $\tau(X, Y)$ qui est la plus fine de toutes). Alors $f$ *est sous-différentiable au point* $x_0$ (cf. par exemple Minty [24]) et de ce qu'on verra au § 7 il résulte que le sous-différentiel $\partial f(x_0)$ est un ensemble (convexe, non vide) *compact* pour la topologie faible $\sigma(Y, X)$. En outre (Moreau [35], [37]), pour tout $u \in X$, la fonction considérée en (5.1) *possède une dérivée à droite de zéro* égale a

$$\frac{d}{dt} f(x_0+tu)\,|_{t=0^+} = \max\{\langle u,y\rangle : y \in \partial f(x_0)\}\,. \tag{5.2}$$

On voit donc que *si $f$, convexe, est finie et continue au point $x_0$ et si le sous-différentiel $\partial f(x_0)$ se réduit à un seul élément, alors $f$ est différentiable au sens de Gâteaux au point $x_0$.* En vertu de (3.5), une condition suffisante pour que $\partial f(x)$ contienne au plus un élément est la *stricte convexité* de $f^*$ sur l'espace $Y$.

Si l'espace $X$ est de dimension finie, on trouvera d'autres précisions sur les *dérivées directionnelles* telles que (5.2) dans Fenchel [18], Rockafellar [54] et, pour les espaces de dimension quelconque, dans Rockafellar [51].

## 6. Sous-différentiel d'une somme.

**6.a.** Soient $h_1$ et $h_2$ dans $\bar{\boldsymbol{R}}^X$. Il est clair que *si* chacune de ces deux fonctions possède une minorante affine continue exacte au point $x$, la somme de ces deux minorantes est une minorante affine continue de $h_1+h_2$, encore exacte au point $x$. On a donc l'inclusion

$$\partial h_1(x)+\partial h_2(x) \subset \partial(h_1+h_2)(x)\,. \tag{6.1}$$

Il est très important pour la pratique d'énoncer des cas où l'inclusion (6.1) se précise en une égalité d'ensembles:

*Si les fonctions $h_1$ et $h_2$ sont convexes et s'il existe un point $x_0$ où les deux fonctions sont finies, l'une d'entre elles continue en ce point, alors pour tout $x \in X$ on a*

$$\partial h_1(x)+\partial h_2(x) = \partial(h_1+h_2)(x)\,. \tag{6.2}$$

Ce résultat est établi par Rockafellar [49] en utilisant les épigraphes; on peut aussi le déduire du résultat obtenu antérieurement par Moreau [35], [37], pour le cas particulier où $h_1$ et $h_2$ appartiennent à $\Gamma_0(X)$.

**6.b.** Exemple. Soit $C$ une partie de $X$ et soit $h$ une fonction définie sur $X$, à valeurs dans $]-\infty, +\infty]$. Le problème très usuel suivant (programmation, théorie de l'approximation, etc. ...): *caractériser, s'il existe, un point $x_1$ en lequel la restriction de $h$ à $C$ atteigne un minimum* équivaut visiblement à: *caractériser un point $x_1$ où la fonction $h+\psi_C$* (partout définie sur $X$) *atteigne un minimum* (cf. §§ 2a, 2b). Cela s'exprime encore par

$$0 \in \partial(h+\psi_C)(x_1)\,. \tag{6.3}$$

Supposons que $h$ soit une *fonction convexe* et $C$ un *ensemble convexe*. Alors les conditions du § 6.a se trouvent réalisées:

— avec $h_1 = h$, $h_2 = \psi_C$, s'il existe un point $x_0 \in C$ où la fonction $h$ est continue,

— avec $h_1 = \psi_C$, $h_2 = h$, s'il existe un point $x_0$ *intérieur à* $C$ où la fonction $h$ est finie.

Dans l'un et l'autre cas on a l'égalité (6.2) partout de sorte que (6.3) fournit la caractérisation suivante du point $x_1$:

— ou bien $x_1 \in C$ confère à $h$ son minimum sur $X$,

— ou bien $C$ est contenu dans un *demi-espace fermé d'appui* de la forme

$$\{x \in X : \langle x - x_1, y_1 \rangle \geqq 0\} \text{ avec } x_1 \in C \text{ et } y_1 \in \partial h(x_1) .$$

Des applications plus raffinées de ce théorème d'addition des sous-différentiels se trouvent dans Rockafellar [49]: on peut notamment en déduire la théorie des multiplicateurs de Lagrange.

## 7. Critères de continuité.

**7.a.** Les énoncés des §§ 5 et 6 attirent l'intérêt sur les cas de continuité d'une fonction convexe.

Supposons d'abord que l'espace $X$ soit de *dimension finie*.

Il est classique que toute fonction $h$ convexe sur $X$, à valeurs dans $]-\infty, +\infty]$, est continue en chaque point *intérieur* de son *domaine effectif*:

$$\operatorname{dom} h = \{x \in X : h(x) < +\infty\} .$$

Elle est donc sous-différentiable en chacun de ces points. Plus généralement, puisque $\operatorname{dom} h$ est une partie convexe d'un espace de dimension finie, cet ensemble possède, comme il est classique, un *intérieur relatif* non vide (c'est à dire l'intérieur de l'ensemble relativement à la topologie de la plus petite variété linéaire le contenant). En tout point de cet intérieur relatif la sous-différentiabilité de $h$ est assurée.

**7.b.** Si $X$ est de dimension infinie, par contre, une fonction convexe, même partout finie, peut n'être continue en aucun point (exemple: une forme linéaire non continue).

Rappelons d'abord (cf. Bourbaki [11]) que si $h$ est une fonction convexe à valeurs dans $]-\infty, +\infty]$, et s'il existe un ouvert $\Omega$ sur lequel $h$ est majorée, alors $h$ est continue sur l'intérieur de $\operatorname{dom} h$. (C'est la conclusion qu'on obtient, par exemple, dès qu'on possède un point où $h$ est semi-continue supérieurement).

On peut en tirer (Rockafellar [50]) que si $X$ est un espace *tonnelé* (par exemple, un espace de Banach) une fonction $f \in \Gamma_0(X)$ est continue en un point $x_0$ si et seulement si $x_0$ est un *point interne* de dom$f$ (c'est à dire que toute droite passant par $x_0$ rencontre dom$f$ selon un segment sur lequel $x_0$ est intérieur).

**7.c.** Reprenant les espaces $X$ et $Y$ mis en dualité par la forme bilinéaire $\langle .,. \rangle$, considérons sur $X$ une topologie $\mathcal{T}$ *non nécessairement compatible avec la dualité.* On a alors (cf. Moreau [37], [40]):

*La fonction $f \in \Gamma_0(X)$ est continue à l'origine de $X$ pour la topologie $\mathcal{T}$ si et seulement si pour tout $r \in \boldsymbol{R}$ l'ensemble*

$$g^{\leqq}(r) = \{y \in Y : f^*(y) \leqq r\}$$

*est une partie de $Y$ équicontinue vis à vis de la topologie $\mathcal{T}$.*

Il suffit d'ailleurs pour cela qu'il existe un $r_0 > \inf f^*$ tel que $g^{\leqq}(r_0)$ soit équicontinu.

Par la règle de translation du § 3.d on caractérise de manière analogue la continuité de $f$ en un point quelconque.

Ce résultat donne de nombreux corollaires utiles, par exemple:

Il existe une topologie (localement convexe séparée) compatible avec la dualité et rendant $f$ continue à l'origine, avec $f(0) < +\infty$ (autrement dit: $f$ est finie continue à l'origine pour la topologie de Mackey $\tau(X,Y)$) si et seulement si les ensembles $g^{\leqq}(r)$ sont *compacts* pour la topologie faible $\sigma(Y,X)$ (Moreau [37]).

De même $f \in \Gamma_0(X)$ est continue à l'origine de $X$ pour la *topologie forte* $\beta(X,Y)$ si et seulement si les ensembles $g^{\leqq}(r)$ sont *bornés* dans $Y$ (cf. § 1.c). On peut en tirer (cf. Rockafellar [50]) que $f \in \Gamma_0(X)$ est finie et continue pour la topologie forte en un point $x_0 \in X$ si et seulement si $x_0$ est point interne de dom$f$.

**7.d.** L'équivalence entre la $\tau$-continuité de $f \in \Gamma_0(X)$ et la compacité faible des « sections » de sa fonction duale a conduit Aggeri [1] à invoquer pour l'étude des fonctions convexes continues, le théorème classique de Krein–Milman sur les *points extrémaux* des ensembles convexes compacts et à généraliser un peu ce théorème (notion de point extrémal d'une fonction numérique). On aboutit ainsi à formuler une proposition duale du théorème de Krein–Milman.

## 8. Assise d'une fonction.

**8.a.** Nous appelons *assise* d'une fonction $h \in \bar{\boldsymbol{R}}^E$ l'ensemble des points où cette fonction est sous-différentiable, c'est à dire

$$A = \{x \in E : \partial h(x) \neq \emptyset\} \subset \operatorname{dom} h .$$

Il résulte du § 6.a que si $h$ est convexe et si $g$ est une fonction convexe partout finie et continue, la somme $g+h$ a même assise que $h$.

Brøndsted et Rockafellar [13] ont amélioré et étendu au cas d'espaces de Banach quelconques des résultats obtenus d'abord par l'auteur pour les espaces de Banach réflexifs (Moreau [35]):

*Si $E$ est un espace de Banach quelconque, toute fonction $f \in \Gamma_0(E)$ est égale à l'enveloppe supérieure de ses minorantes affines continues exactes* (§ 2.a). Ce résultat découle par dualité du suivant: *Si $E$ est un espace de Banach quelconque, toute $f \in \Gamma_0(E)$ est déterminée par les valeurs qu'elle prend sur son assise $A$; on a, pour tout $x \in E$,*

$$f(x) = \liminf_{u \to x} [f(u) + \psi_A(u)] .$$

En d'autres termes, *$f$ est la régularisée semi-continue inférieurement de sa restriction à son assise.* Il en résulte notamment que $A$ est dense dans $\operatorname{dom} f$; plus généralement, si on considère un nombre $r \in \overline{\boldsymbol{R}}$ et l'ensemble convexe

$$\{x \in E : f(x) < r\}$$

les points de cet ensemble tels que $\partial f \neq \emptyset$ y sont denses.

Ces propriétés ne sont plus vraies en général pour un espace normé quelconque: on peut en effet, dans des espaces normés convenables, construire des fonctions appartenant à $\Gamma_0$ et dont l'assise est vide.

**8.b.** Les résultats précédents sont dans la ligne des travaux de Bishop et Phelps [10], Phelps [45], sur la « subréflexivité » des espaces de Banach quelconques et sur les hyperplans d'appuis des ensembles convexes fermés de ces espaces.

Dans les raisonnements de Brøndsted et Rockafellar [13], Rockafellar [51], la notion de *sous-gradient à $\varepsilon$ près* joue un rôle essentiel: reprenant les espaces en dualité $X$ et $Y$, on posera, par analogie avec (3.5), la notation

$$\partial_\varepsilon f(x) = \{y \in Y : f(x) + f^*(y) - \langle x, y \rangle \leqq \varepsilon\} .$$

Cet ensemble est convexe et fermé dans $Y$, non vide si $f \in \Gamma_0(X)$ et $f(x) < +\infty$. Le lemme fondamental de Brøndsted et Rockafellar s'énonce alors:

*Supposons que $X$ soit un espace de Banach et $f \in \Gamma_0(X)$. Soit $\varepsilon > 0$ et soient $x \in X$, $y \in Y$ tels que $y \in \partial_\varepsilon f(x)$; choisissons $\lambda \in ]0, +\infty[$.*

*Alors il existe $\bar{x} \in X$ et $\bar{y} \in Y$ tels que*

$$\|\bar{x}-x\| \leqq \lambda, \quad \|\bar{y}-y\| \leqq \frac{\varepsilon}{\lambda}, \quad \bar{y} \in \partial f(\bar{x}).$$

**9. Applications sous-différentielles.**

**9.a.** Considérons encore les espaces en dualité $X$ et $Y$. Soit $h \in \bar{\boldsymbol{R}}^X$. En associant à tout $x \in X$ le sous-différentiel $\partial h(x)$, lequel est une partie (éventuellement vide) de $Y$, on définit une *application multivoque* de $X$ dans $Y$.

Donnons d'abord une propriété de continuité locale (Moreau [40]):

*Munissons $X$ d'une topologie compatible avec la dualité et $Y$ de la topologie faible $\sigma(Y,X)$. Si la fonction numérique $h$ est continue au point $x_0 \in X$, l'application multivoque $x \to \partial h(x)$ est semi-continue supérieurement en ce point, c'est à dire que, quel que soit $\Omega$, ouvert dans $Y$ contenant $\partial h(x_0)$, il existe un voisinage $V$ de $x_0$ dans $X$ tel que*

$$x \in V \Rightarrow \partial h(x) \subset \Omega.$$

En outre, pour tout $x \in V$, l'ensemble $\partial h(x)$ est alors équicontinu, donc faiblement compact.

On peut d'ailleurs améliorer ce résultat en remplaçant la topologie $\sigma(Y,X)$ par la topologie de la convergence uniforme dans les parties compactes de $X$.

Naturellement, si pour tout $x$ dans un voisinage de $x_0$, l'ensemble $\partial h(x)$ est constitué d'un seul élément, on trouve la continuité au sens habituel de cette application univoque. Tel est le cas, en particulier, si $h$ (continue) est convexe et faiblement différentiable (cf. § 5.b).

**9.b.** Soit, d'une manière générale, $\Phi$ une application multivoque associant à tout $x \in X$ une partie $\Phi(x)$ (éventuellement vide) de $Y$. Avec G. J. Minty on dit que $\Phi$ est *monotone* si, quels que soient $x$ et $x'$ dans $X$, quels que soient $y \in \Phi(x)$ et $y' \in \Phi(x')$ on a

$$\langle x-x', \ y-y' \rangle \geqq 0.$$

(C'est là, en fait, une propriété du graphe de $\Phi$ dans l'espace produit $X \times Y$ et les deux espaces jouent des rôles symétriques: l'application inverse $\Phi^-$ est aussi bien monotone.)

$\Phi$ est dite *monotone maximale* si elle n'admet pas d'extension stricte qui soit encore monotone: autrement dit en ce cas, si un couple $a \in X$, $b \in Y$ est tel que l'on ait

$$\langle a-x, \ b-y \rangle \geqq 0$$

pour tout $x \in X$ et tout $y \in \Phi(x)$, alors $b \in \Phi(a)$.

L'intérêt principal de la maximalité est de fournir des *théorèmes d'existence* pour certains problèmes fonctionnels (Minty [23]).

**9.c.** Soit une fonction quelconque $h \in \overline{\boldsymbol{R}}^X$. On vérifie immédiatement que l'application multivoque $x \to \partial h(x)$ est *monotone*. Si $h$ est *convexe*, partout finie et continue sur $X$, Minty [24] montre que cette application est monotone maximale.

Brøndsted et Rockafellar [13] ont montré que, si $X$ est un *espace de Banach*, quelle que soit $f \in \Gamma_0(X)$, l'application $\partial f$ est monotone maximale. Dans le cas particulier d'un espace de Hilbert, le fait résulte d'ailleurs très simplement de la théorie des applications « prox » développée par l'auteur (Moreau [41]). Pour un espace normé quelconque, par contre, l'exemple d'une fonction $f \in \Gamma_0$ telle que $\partial f$ soit partout vide montre que la maximalité de l'application monotone $\partial f$ n'est plus assurée en général.

**9.d.** On peut d'autre part se demander dans quelle mesure une fonction $f \in \Gamma_0(X)$ est déterminée par son application sous-différentielle. Rockafellar [51] a montré que si $X$ est un espace de Banach et si $f$ et $g$ appartiennent à $\Gamma_0(X)$ l'inclusion

$$\partial f(x) \subset \partial g(x) \quad \text{pour tout } x \in X$$

implique

$$g = f + \text{const.}$$

Le même contre-exemple que ci-dessus montre que cette propriété n'est plus vraie en général dans un espace normé quelconque. (Pour le cas particulier d'un espace de Hilbert, la théorie des applications « prox » fournit, là encore, une démonstration très simple.)

**9.e.** Sauf dans le cas particulier $X = \boldsymbol{R}$, il existe des applications monotones, même maximales, qui ne sont pas des sous-différentielles. Pour caractériser les sous-différentielles, Rockafellar [51] introduit la notion suivante: une application multivoque $\Phi$ de $X$ dans $Y$ est dite *cycliquement monotone* si, pour toute famille finie de couples $(x_i, y_i) \in X \times Y$, $i = 0, 1, \ldots, n$, tels que

$$y_i \in \Phi(x_i)$$

on a

$$\langle x_0 - x_1, y_0 \rangle + \langle x_1 - x_2, y_1 \rangle + \ldots + \langle x_n - x_0, y_n \rangle \geqq 0$$

(en particulier pour $n = 1$, cela donne l'inégalité de monotonie de Minty).

Il montre que cette propriété est nécessaire et suffisante pour l'existence d'une fonction $f \in \Gamma_0(X)$ telle que, pour tout $x \in X$, on ait

$$\Phi(x) \subset \partial f(x) .$$

**9.f.** Aggeri et Lescarret [2] ont constaté que les « duality mappings » d'un espace de Banach dans son dual, considérées par Beurling et Livingston [8], Browder [14], [15], étaient précisément des applications sous-différentielles. Plus généralement, ils partent de deux espaces en dualité $X$ et $Y$ et d'un couple d'ensembles $A \subset X$ et $B \subset Y$ *mutuellement polaires*. La fonction $a$, *jauge* de $A$, est alors la fonction d'appui de $B$ et la *jauge* $b$ de $B$ est la fonction d'appui de $A$. Ils considèrent d'autre part un couple de fonctions, $f$ et $g$, définies sur $[0, +\infty]$, qui sont *conjuguées de Young*, c'est à dire que

$$f(\xi) = \sup\{\xi\eta - g(\eta) : \eta \in [0, +\infty[\}$$

et égalité analogue en échangeant $f$ et $g$ ainsi que $\xi$ et $\eta$. Ils montrent alors que les fonctions $F = f \circ a \in \Gamma_0(X)$ et $G = g \circ b \in \Gamma_0(Y)$ sont des fonctions duales et étudient l'application $\partial F$. On retrouve une « duality mapping » de Beurling–Livingston–Browder en prenant pour $A$ et $B$ les boules unité d'un espace de Banach réflexif et de son dual et en particularisant $f$.

## BIBLIOGRAPHIE


1. J. C. Aggeri, *Les fonctions convexes continues et le théorème de Krein–Milman*, C. R. Acad. Sci. Paris 262 (1966), 229–232.
2. J. C. Aggeri et C. Lescarret, *Fonctions convexes duales associées à un couple d'ensembles mutuellement polaires*, C. R. Acad. Sci. Paris 260 (1965), 6011–6014.
3. R. E. Bellman and W. Karush, *On a new functional transform in analysis; the maximum transform*, Bull. Amer. Math. Soc. 67 (1961), 501–503.
4. R. E. Bellman and W. Karush, *Mathematical Programming and the Maximum Transform*, J. Soc. Indust. Appl. Math. 1 (1962), 550–567.
5. R. E. Bellman and W. Karush, *On the Maximum Transform and Semigroups of Transformations*, Bull. Amer. Math. Soc. 68 (1962), 516–518.
6. R. E. Bellman and W. Karush, *On the Maximum Transform*, J. Math. Anal. Appl. 6 (1963), 67–74.
7. R. E. Bellman and W. Karush, *Functional Equations and the Theory of Dynamic Programming; An Application of the Maximum Transform*, J. Math. Anal. Appl. 6 (1963), 155–157.
8. A. Beurling and A. E. Livingston, *A theorem on duality mappings in Banach spaces*, Ark. Mat. 4 (1960–63), 405–411.
9. Z. Birnbaum und W. Orlicz, *Über die Verallgemeinerung des Begriffes der zueinander konjugierten Potenzen*, Studia Math. 3 (1931), 1–67.

10. E. Bishop and R. R. Phelps, *The support functionals of a convex set*, Proceedings of Symposia in Pure Mathematics vol. 7, Amer. Math. Soc. 1963, 27–35.
11. N. Bourbaki, *Espaces vectoriels topologiques*, Hermann, Paris, 1953–55.
12. A. Brøndsted, *Conjugate convex functions in topological vector spaces*, Mat. Fys. Medd. Dan. Vid. Selsk. 34, n° 2 (1964).
13. A. Brøndsted and R. T. Rockafellar, *On the subdifferentiability of convex functions*, Proc. Amer. Math. Soc. 16 (1965), 605–611.
14. F. E. Browder, *On a theorem of Beurling and Livingstone*, Canad. J. Math. 17 (1965), 367–372.
15. F. E. Browder, *Multivalued monotone non-linear mappings and duality mappings in Banach spaces*, Trans. Amer. Math. Soc. 118 (1965), 338–351.
16. U. Dieter, *Theorie der Optimierungsaufgaben in topologischen Vektorräumen*, Habilitationschrift, Christian-Albrechts-Universität, Kiel, 1964. *Optimierungsaufgaben in topologischen Vektorräumen* I. Z. Wahrscheinlichkeitstheorie verw. Gebiete 5 (1966), 89–117.
17. W. Fenchel, *On conjugate convex functions*, Canad. J. Math. 1 (1949), 73–77.
18. W. Fenchel, *Convex cones, sets, and functions*, Lecture notes, Princeton University, 1953.
19. W. L. Jones, *On conjugate functionals*, Doctoral dissertation, Columbia University, 1960.
20. S. Karlin, *Mathematical Methods and Theory in Games, Programming and Economics*, vol. 1, Addison Wesley, Reading (Mass.) 1960.
21. C. Lescarret, *Cas d'addition des applications monotones maximales dans un espace de Hilbert*, C. R. Acad. Sci. Paris 261 (1965), 1160–1163.
22. S. Mandelbrojt, *Sur les fonctions convexes*, C. R. Acad. Sci. Paris 209 (1939), 977–978.
23. G. J. Minty, *Monotone (non linear) operators in Hilbert space*, Duke Math. J. 29 (1962), 341–346.
24. G. J. Minty, *On the monotonicity of the gradient of a convex function*, Pac. J. Math. 14 (1964), 243–247.
25. J. J. Moreau, *Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires*, C. R. Acad. Sci. Paris, 255 (1962), 238–240.
26. J. J. Moreau, *Fonctions convexes en dualité*, Faculté des Sciences de Montpellier, Séminaires de Mathématiques, 1962 (multigraphié, 18 pages).
27. J. J. Moreau, *Fonctions convexes duales et points proximaux dans un espace hilbertien*, C. R. Acad. Sci. Paris 255 (1962), 2897–2899.
28. J. J. Moreau, *Applications «prox»*, Faculté des Sciences de Montpellier, Séminaires de Mathématiques, 1963 (multigraphié, 20 pages).
29. J. J. Moreau, *Les liaisons unilatérales et le principe de Gauss*, C. R. Acad. Sci. Paris 256 (1963), 871–874.
30. J. J. Moreau, *Propriétés des applications «prox»*, C. R. Acad. Sci. Paris 256 (1963), 1069–1071.
31. J. J. Moreau, *Inf-convolution*, Faculté des Sciences de Montpellier, Séminaires de Mathématiques, 1963 (multigraphié, 48 pages).
32. J. J. Moreau, *Inf-convolution des fonctions numériques sur un espace vectoriel*, C. R. Acad. Sci. Paris 256 (1963), 5047–5049.
33. J. J. Moreau, *Fonctions à valeurs dans* $[-\infty, +\infty]$; *notions algébriques*, Faculté des Sciences de Montpellier, Séminaires de Mathématiques, 1963 (multigraphié, V+37 pages).
34. J. J. Moreau, *Remarques sur les fonctions à valeurs dans* $[-\infty, +\infty]$ *définies sur un demi-groupe*, C. R. Acad. Sci. Paris 257 (1963), 3107–3109.

35. J. J. Moreau, *Etude locale d'une fonctionnelle convexe*, Faculté des Sciences de Montpellier, Séminaires de Mathematiques, 1963 (multigraphié, 25 pages).

36. J. J. Moreau, *Fonctionnelles sous-différentiables*, C. R. Acad. Sci. Paris 257 (1963), 4117–4119.

37. J. J. Moreau, *Sur la fonction polaire d'une fonction semi-continue supérieurement*, C. R. Acad. Sci. Paris 258 (1964), 1128–1131.

38. J. J. Moreau, *Théorème «inf-sup»*, C. R. Acad. Sci. Paris 258 (1964), gr. 1, 2720–2722.

39. J. J. Moreau, *Sur la naissance de la cavitation dans une conduite*, C. R. Acad. Sci. Paris 259 (1964), gr. 2, 3948–3951.

40. J. J. Moreau, *Semi-continuité du sous-gradient d'une fonctionnelle*, C. R. Acad. Sci. Paris 260 (1965), gr. 1, 1067–1070.

41. J. J. Moreau, *Proximité et dualité dans un espace hilbertien*, Bull. Soc. Math. France 93 (1965), 273–299.

42. J. J. Moreau, *One-sided constraints in hydrodynamics*, à paraître dans J. Abadie, editor, *Non linear programming; a course*, North Holland Pub. Co.

43. J. J. Moreau, *Quadratic programming in Mechanics: Dynamics of one-sided constraints*, J. Soc. Indust. Appl. Math. Ser. A Control 4 (1966), 153–158.

44. J. J. Moreau, *Convexity and duality*, à paraître dans *Functional analysis and optimization*, Academic Press.

45. R. R. Phelps, *Weak* support points of convex sets in $E^*$*, Israel J. Math. 2 (1964), 177–182.

46. R. T. Rockafellar, *Convex functions and dual extremum problems*, Thesis, Harvard University, 1963.

47. R. T. Rockafellar, *Duality theorems for convex functions*, Bull. Amer. Math. Soc. 70 (1964), 189–192.

48. R. T. Rockafellar, *Minimax theorems and conjugate saddle-functions*, Math. Scand. 14 (1964), 151–173.

49. R. T. Rockafellar, *Extension of Fenchel's duality theorem for convex functions*, Duke Math. J. 33 (1966), 81–90.

50. R. T. Rockafellar, *Level sets and Continuity of conjugate convex functions*, Trans. Amer. Math. Soc. 123 (1966), 46–63.

51. R. T. Rockafellar, *Characterization of the subdifferentials of convex functions*, Pacific J. Math. 17 (1966), 497–510.

52. R. T. Rockafellar, *Dual extremum problems involving convex functions*, à paraître dans Pacific J. Math.

53. R. T. Rockafellar, *A general correspondance between dual minimax problems and convex programs*, à paraître.

54. R. T. Rockafellar, *Conjugates and Legendre Transforms of convex functions*, Canad. J. Math. 19 (1967), 200–205.