Journal of Global Optimization **12:** 373–382, 1998.

On the Minimization of Difference Functions

JEAN-PAUL PENOT
Université de Pau, Laboratoire de Mathématiques Appliquées, URA-CNRS 1204, Avenue de l'Université, F-64000 Pau-France

(Received 28 February 1996; accepted 2 September 1997)

Abstract. Simple necessary optimality conditions are formulated for a function f of the form $f = g - h$, where g and h are nonsmooth functions. Related sufficient conditions are given for local minimization and global minimization.

Key words: Critical point, d.c. Functions, Difference of convex functions, Generalized derivatives, Minimization, Optimality conditions, Palais-Smale condition, Subdifferentials.

1. Introduction

There is an abundance of problems which involve difference functions. We call a function $f : X \to \overline{\mathbb{R}}$ on a normed vector space (n.v.s.) a difference function if it has a decomposition of the form

$$f = g - h \tag{1}$$

with $g : X \to \overline{\mathbb{R}}, h : X \to \mathbb{R}$. Although there exist tools for handling sums and differences of extended reals (see Moreau [10]) we limit our study to the present case for the sake of simplicity. Such functions are of interest when g is convex and h is differentiable [9, 22] or when g and h are both convex. In that case one says that f is a d.c. function.

For instance, if one maximizes a convex function h over a convex subset C, one may introduce $f = g - h$, where $g = i_C$ is the indicator function of $C(i_C(x) = 0$ for $x \in C, i_C(x) = \infty$ otherwise), and minimize f.

Also, it has been observed by Asplund [1] that the square d_C^2 of the distance function d_C to an arbitrary closed subset C of a Hilbert space, given by $d_C(x) = \inf_{y \in C} \|x - y\|$, is a d.c. function

$$\tfrac{1}{2} d_C^2(x) = \tfrac{1}{2}\|x\|^2 - \left(i_C + \tfrac{1}{2}\| \cdot \|^2\right)^* (x),$$

where i_C is the indicator function of C (given by $i_C(x) = 0$ if $x \in C, +\infty$ otherwise) and f^* is the convex conjugate of f.

Such functions also appear in the study of favorable classes [3, 15, 19].

Let us note that one disposes of duality results for the class of d.c. functions which make this class very attractive (see, for example, [6, 18, 23–25, 27 and 28]).

Moreover, efficient algorithms have been devised for problems involving such functions (see [24, 25 and 29] and references therein).

Here we do not impose a priori convexity or differentiability assumptions on g and h. Then it may be objected that such a decomposition is spurious. Still the structure of the problem at hand may strongly suggest to take into account a decomposition as in (1). This is the case if, for instance, g and h are suprema of finite families of functions of class C^1. More generally, it is the case when both g and h are *tangentially convex,* i.e. when their lower (or contingent) directional derivatives are convex. Another case in which such a decomposition appears in a natural way is when f is deduced from a given bifunction $l : X \times Y \to \mathbb{R}$ by $f(x, y) := l(x, y_0) - l(x_0, y)$, so that the point $(x_0, y_0) \in X \times Y$ is a saddle point of l iff it is a minimizer of f (with value 0).

The present paper focuses on optimality conditions.

For this aim we use simple and classical concepts of nonsmooth analysis, i.e., contingent and Fréchet subdifferentials which coincide in finite dimensions and which are among the most basic tools of nonsmooth analysis. As they do not enjoy a rich calculus, they are often considered as non proper tools. Let us note however that in a large class of Banach spaces they satisfy fuzzy calculus rules [5, 8, 11] and a Mean Value Theorem (see [16] and references therein). It is the purpose of this note to show that these rough subdifferentials may surprisingly serve to devise simple and useful optimality conditions for difference functions. Let us observe that the optimality conditions we give here are local, even when g and h are convex. Thus, they can be used as a first step paving the way for global conditions ([7], for instance) for which concrete testing may be involved or costly.

We note that under generalized convexity properties these conditions become global conditions.

2. First order conditions

In the sequel X is a normed vector space with dual X^*. Its closed unit ball is denoted by B and the closed unit ball of X^* is denoted by B^*. Recall that $\overline{x}^* \in X^*$ is an element of the *firm* (or Fréchet) *subdifferential* $\partial^- f(\overline{x})$ of a function f at $\overline{x} \in X$ if for each $\varepsilon > 0$ there exists $\delta > 0$ such that for each x in the ball $B(\overline{x}, \delta)$ with center $\overline{x}$ and radius δ one has

$$f(x) \geqslant f(\overline{x}) + \langle \overline{x}^*, x - \overline{x} \rangle - \varepsilon \|x - \overline{x}\|. \tag{2}$$

Our first result is an immediate consequence of this definition. It coincides with a well-known result when g and h are convex since then $\partial^- h(x)$ and $\partial^- g(x)$ are the subdifferentials of convex analysis.

PROPOSITION 2.1. *If $\overline{x}$ is a local minimizer of $f = g - h$ then*

$$\partial^- h(\overline{x}) \subset \partial^- g(\overline{x}).$$

Proof. For some $\rho > 0$ and each x in the ball $B(\overline{x}, \rho)$ with center $\overline{x}$ and radius ρ we have $f(x) \geqslant f(\overline{x})$ hence,

$$\overline{g}(x) := g(x) + h(\overline{x}) \geqslant \overline{h}(x) := h(x) + g(\overline{x})$$

and $\overline{g}(\overline{x}) = \overline{h}(\overline{x})$. As $\partial^- h(\overline{x}) = \partial^- \overline{h}(\overline{x})$ and $\partial^- g(\overline{x}) = \partial^- \overline{g}(\overline{x})$, the result follows from the obvious inclusion $\partial^- \overline{h}(\overline{x}) \subset \partial^- \overline{g}(\overline{x})$. □

The following conditions can be proved similarly (when g and h are convex, see [23]). They are formulated in terms of the *contingent derivative* given by

$$h'(\overline{x}, v) = \liminf_{t \downarrow 0, u \to v} \frac{1}{t} (h(\overline{x} + tu) - h(\overline{x})) \tag{3}$$

and of the *contingent subdifferential* defined by

$$\partial h(\overline{x}) := \{\overline{x}^* \in X^* : \overline{x}^* \leqslant h'(\overline{x}, .)\}. \tag{4}$$

PROPOSITION 2.2. *If $\overline{x}$ is a local minimizer of $f = g - h$ then*

$$h'(\overline{x}, v) \leqslant g'(\overline{x}, v) \text{ for each } v \in X, \tag{5}$$

$$\partial h(\overline{x}) \subset \partial g(\overline{x}). \tag{6}$$

The preceding inclusions can be deduced from the obvious general necessary condition $0 \in \partial f(\overline{x})$ (respectively $0 \in \partial^- f(\overline{x})$). It suffices to observe that $g = f + h$, so that one has $\partial f(\overline{x}) + \partial h(\overline{x}) \subset \partial g(\overline{x})$ (and a similar inclusion with the Fréchet subdifferential), hence,

$$\partial f(\overline{x}) \subset \partial g(\overline{x}) \overset{*}{-} \partial h(\overline{x}), \tag{7}$$

where for two subsets A, B of X^* one denotes by

$$A \overset{*}{-} B := \{x^* \in X^* : x^* + B \subset A\} \tag{8}$$

their star difference.

Analogous inequalities hold for the *incident derivative* (or upper epi-derivative) h^i given by

$$h^i(\overline{x}, v) = \sup_{\varepsilon > 0} \limsup_{t \downarrow 0} \inf_{u \in B(v, \varepsilon)} \frac{1}{t} (h(\overline{x} + tu) - h(\overline{x})),$$

for the *lower hypoderivative* given by

$$h^\natural(\overline{x}, v) = -(-h)^i(\overline{x}, v)$$

and the *upper hypoderivative* given by

$$h^\sharp(\overline{x}, v) = -(-h)'(x, v) = \lim_{t \downarrow 0} \sup_{u \to v} \frac{1}{t} (h(\overline{x} = tu) - h(\overline{x})).$$

Similarly, one can use the *lower and upper radial derivatives* given respectively by

$$h'_r(\overline{x}, v) = \liminf_{t \downarrow 0} \frac{1}{t}\,(h(\overline{x} + tv) - h(\overline{x}))$$

$$h^\sharp_r(\overline{x}, v) = -(-h)'_r(\overline{x}, v) = \limsup_{t \downarrow 0} \frac{1}{t}\,(h(\overline{x} + tv) - h(\overline{x})).$$

More generally, if $f^?$ denotes a derivative of f of some sort which is homotone in the sense that $g^?(\overline{x}, v) \geqslant h^?(\overline{x}, v)$ whenever $g \geqslant h$ and $g(\overline{x}) = h(\overline{x})$, then one has the necessary condition $g^?(\overline{x}, v) \geqslant h^?(\overline{x}, v)$ for each v in X and the related inclusion for the subdifferential $\partial^?$ associated with this derivative.

In the sufficient condition which follows one says that h is *semi-differentiable* at $\overline{x}$ if $h'(\overline{x}, v) = h^\sharp(\overline{x}, v)$ for each $v \in X$.

PROPOSITION 2.3. *Suppose X is finite dimensional, f is finite at $\overline{x}$ and h is semi-differentiable at $\overline{x}$. If*

$$h'(\overline{x}, v) < g'(\overline{x}, v) \text{ for each } v \in X\backslash\{0\}, \tag{9}$$

then $\overline{x}$ is a local strict minimizer of f. The same conclusion holds if $h'(\overline{x}, .)$ is convex and finite and if

$$\partial h(\overline{x}) \subset \mathit{int}\, \partial g(\overline{x}). \tag{10}$$

Proof. Suppose on the contrary that there exists a sequence (x_n) with limit $\overline{x}$ such that $x_n \neq \overline{x}$ and $f(x_n) \leqslant f(\overline{x})$ for each n. Let us set $t_n = \|x_n - \overline{x}\|$, $v_n = t_n^{-1}(x_n - \overline{x})$. Without loss of generality we may suppose that (v_n) converges to some unit vector v. As for each n

$$g(\overline{x} + t_n v_n) - g(\overline{x}) \leqslant h(\overline{x} + t_n v_n) - h(\overline{x}),$$

dividing by t_n and taking limits we get

$$g'(\overline{x}, v) \leqslant h^\sharp(\overline{x}, v),$$

which is in contradiction with our assumption. If (10) holds and $h'(\overline{x}, .)$ is convex and finite, hence continuous, for each $v \in X\backslash\{0\}$ we can find $x^* \in \partial h(\overline{x})$ such that $\langle x^*, v\rangle = h'(\overline{x}, v)$. As there exists $\varepsilon > 0$ with $x^* + \varepsilon B^* \subset \partial g(\overline{x})$ we have

$$g'(\overline{x}, v) \geqslant \sup_{u^* \in B^*} \langle x^* + \varepsilon u^*, v\rangle = h'(\overline{x}, v) + \varepsilon\|v\|$$

and we are back to our first assumption. □

COROLLARY 2.4. *Suppose X is finite dimensional, f is finite at $\overline{x}$ and h is convex. If*

$$\partial h(\overline{x}) \subset \mathit{int}\, \partial g(\overline{x})$$

then $\overline{x}$ is a local strict minimizer of f.

Proof. As h is finite, by our standing assumption, and convex, it is locally Lipschitzian, so that $h'(\overline{x}, .)$ is convex and finite and the preceding proposition applies since h is semi-differentiable at $\overline{x}$. □

The assumptions of the preceding proposition have to be reinforced in the infinite dimensional case. Let us say that the function g has a *firm lower derivative* at $\overline{x}$ if

$$\liminf_{\|v\| \searrow 0} \|v\|^{-1}(g(\overline{x}+v) - g(\overline{x}) - g'(\overline{x}, v)) \geqslant 0.$$

We also say that h has a firm upper derivative at $\overline{x}$ if $-h$ has a firm lower derivative at $\overline{x}$.

PROPOSITION 2.5. *Suppose g (respectively h) has a firm lower (respectively upper) derivative at $\overline{x}$. Then each of the following conditions suffices to ensure that $\overline{x}$ is a strict local minimizer of f:*
(a) there exists some $\varepsilon > 0$ such that $g'(\overline{x}, v) \geqslant h'(\overline{x}, v) + \varepsilon\|v\|$ for each $v \in X$;
(b) $h'(\overline{x}, .)$ is convex and there exists $\varepsilon > 0$ such that

$$\partial h(\overline{x}) + \varepsilon B^* \subset \partial g(\overline{x}). \tag{11}$$

Proof. Let us first observe that Assumption (b) implies Assumption (a) since for each $v \in X$ we have $h'(\overline{x}, v) = \sup_{x^* \in \partial h(\overline{x})} \langle x^*, v\rangle$ under our hypothesis, hence,

$$\begin{aligned} g'(\overline{x}, v) &\geqslant \sup_{x^* \in \partial g(\overline{x})} \langle x^*, v\rangle \geqslant \sup_{x^* \in \partial h(\overline{x})} \langle x^*, v\rangle + \sup_{u^* \in \varepsilon B^*} \langle u^*, v\rangle \\ &= h'(\overline{x}, v) + \varepsilon\|v\|. \end{aligned}$$

Now Assumption (a) ensures the existence of some $\rho > 0$ such that for $v \in \rho B \backslash \{0\}$ one has

$$\begin{aligned} g(\overline{x}+v) &> g(\overline{x}) + g'(\overline{x}, v) - (\varepsilon/2)\|v\|, \\ -h(\overline{x}+v) &> -h(\overline{x}) - h'(\overline{x}, v) + (\varepsilon/2)\|v\|, \end{aligned}$$

and the result follows by adding the respective sides of these relations. □

The conditions of the following global sufficient condition are reminiscent to the notion of invexity to which a number of papers have been devoted during the last few years. When v^* takes its values into $\partial f(\overline{x})$, it reduces to this notion.

PROPOSITION 2.6. *Let $\overline{x} \in X$. Suppose there exist mappings $x \mapsto v(x)$ and $x \mapsto v^*(x)$ from X into X and X^*, respectively, such that*

$$\begin{aligned} g(\overline{x}+x) &\geqslant g(\overline{x}) + \langle v^*(x), v(x)\rangle \\ h(\overline{x}+x) &\leqslant h(\overline{x}) + \langle v^*(x), v(x)\rangle. \end{aligned}$$

Then $\overline{x}$ is a global minimizer of f.

The proof is immediate. Note that for g convex and h concave differentiable one can take $v(x) = x - \overline{x}, v^*(x) = h'(\overline{x})$ when the necessary condition of Proposition 2.2 is satisfied.

3. Second order conditions

Let us introduce some classical second order notions. The *lower parabolic second derivative* of h at $\overline{x}$ in the directions v, w is

$$\ddot{h}(\overline{x}, v, w) = \liminf_{\substack{t \downarrow 0 \\ z \to w}} \frac{2}{t^2} \left(h\left(\overline{x} + tv + \frac{1}{2}\, t^2 z\right) - h(\overline{x}) - h'(\overline{x}, v)\right);$$

this notion, introduced in [12] and [13], is a variant of the genuine parabolic second derivative first given in [2].

The *lower second epi-derivative* of h at $\overline{x}, \overline{x}^*$ is given by

$$h''(\overline{x}, \overline{x}^*, v) = \liminf_{\substack{t \downarrow 0 \\ u \to v}} \frac{2}{t^2} \left(h(\overline{x} + tu) - h(\overline{x}) - \langle \overline{x}^*, tu \rangle \right)$$

(see [20] and [21]). Some relationships between these two derivatives are given in [21], [4] and [13]. The proofs of the following two results are similar to the proofs of Proposition 2.1 and 2.2, respectively. Here, $L_s(X, X^*)$ denotes the space of linear symmetric operators from X into X^*.

PROPOSITION 3.1. *Suppose $\overline{x}$ is a local minimizer of f. Then for each $v \in X$ one has $g'(\overline{x}, v) \geqslant h'(\overline{x}, v)$ and if equality holds one has for each $w \in X$*

$$\ddot{g}(\overline{x}, v, w) \geqslant \ddot{h}(\overline{x}, v, w).$$

Moreover, $\partial h(\overline{x}) \subset \partial g(\overline{x})$ and for each $\overline{x}^* \in \partial h(\overline{x})$ one has

$$g''(\overline{x}, \overline{x}^*, v) \geqslant h''(\overline{x}, \overline{x}^*, v) \text{ for each } v \in X.$$

In the following proposition we use the set $\partial^2 h(\overline{x}, \overline{x}^*)$ of *firm subhessians* of h at $(\overline{x}, \overline{x}^*)$ which is the set of $A \in L_s^2(X, X^*)$ such that

$$\liminf_{\substack{x \to 0 \\ x \neq 0}} \frac{1}{\|x\|^2} \left[h(\overline{x} + x) - h(\overline{x}) - \langle \overline{x}^*, x \rangle - \frac{1}{2} \langle Ax, x \rangle \right] \geqslant 0.$$

(see [14]). This statement is a second order version of Proposition 2.1. Its proof is immediate.

PROPOSITION 3.2. *Suppose $\overline{x}$ is a local minimizer of f. Then for each $\overline{x}^* \in X^*$ the set of firm subhessians of h at $(\overline{x}, \overline{x}^*)$ is contained in the set $\partial^2 g(\overline{x}, \overline{x}^*)$ of firm sub-hessians of g at $(\overline{x}, \overline{x}^*)$.*

A second order sufficient condition can be given along the lines of Proposition 2.5. Here we use the second order upper hypoderivative $h^{\sharp\sharp}$ of h given by $h^{\sharp\sharp} := -(-h)''$ and we say that h is twice semi-differentiable at $(\overline{x}, \overline{x}^*)$ if for each $v \in X$ we have

$$h''(\overline{x}, \overline{x}^*, v) = h^{\sharp\sharp}(\overline{x}, \overline{x}^*, v) := \limsup_{\substack{t \downarrow 0 \\ u \to v}} \frac{2}{t^2} \left(h(\overline{x} + tu) - h(\overline{x}) - \langle \overline{x}^*, tu \rangle\right),$$

or, in other terms, if $2t^{-2}(h(\overline{x} + tu) - h(\overline{x}) - \langle \overline{x}^*, tu \rangle) \to h''(\overline{x}, \overline{x}^*, v)$ as $t \downarrow 0, u \to v$.

PROPOSITION 3.3. *Suppose X is finite dimensional and $g'(\overline{x}, v) \geqslant h'(\overline{x}, v)$ for each $v \in X$. If for each $v \in X \backslash \{0\}$ such that $g'(\overline{x}, v) = h'(\overline{x}, v)$ there exists some $\overline{x}^* \in \partial h(\overline{x})$ such that*

$$g''(\overline{x}, \overline{x}^*, v) > h''(\overline{x}, \overline{x}^*, v),$$

and if h is twice semi-differentiable at $(\overline{x}, \overline{x}^*)$ then $\overline{x}$ is a strict local minimizer of f.

Proof. Suppose again there exists a sequence (x_n) in $X \backslash \{0\}$ with limit $\overline{x}$ such that $f(x_n) \leqslant f(\overline{x})$. Let $x_n = \overline{x} + t_n v_n$ with $t_n = \|x_n - \overline{x}\|$; we may suppose $(v_n) \to v$ for some $v \neq 0$. Then we have $g'(\overline{x}, v) \leqslant h'(\overline{x}, v)$, hence, $g'(\overline{x}, v) = h'(\overline{x}, v)$. Let $\overline{x}^*$ be associated to v such that $g''(\overline{x}, \overline{x}^*, v) > h''(\overline{x}, \overline{x}^*, v)$. Then, as $h''(\overline{x}, \overline{x}^*, v) = h^{\sharp\sharp}(\overline{x}, \overline{x}^*, v)$, for n large enough we have

$$g(\overline{x} + t_n v_n) - g(\overline{x}) > h(\overline{x} + t_n v_n) - h(\overline{x}).$$

which is in contradiction with $f(\overline{x} + t_n v_n) \leqslant f(\overline{x})$. □

Note that, when h is semi-differentiable, our assumptions ensure that $f''(\overline{x}, 0, v)$ is positive for each non null vector v, a sufficient optimality condition, even when one does not know that $0 \in \partial f(\overline{x})$.

In the following obvious result a notion of sinvexity (or second order invexity) is introduced.

Here, a family $(g_i)_{i \in I}$ of functions is said to be simultaneously *sinvex* if there exist mappings $x \mapsto v(x)$, $x \mapsto v^*(x)$, $x \mapsto A(x)$ from X into $X, X^*, L_s(X, X^*)$ respectively such that for each $i \in I$ and for each $x \in X$

$$g_i(\overline{x} + x) \geqslant g_i(\overline{x}) + \langle v^*(x), v(x) \rangle + \tfrac{1}{2} \langle A(x) v(x), v(x) \rangle.$$

Then one says that $(g_i)_{i \in I}$ is sinvex with respect to $v(.), v^*(.), A(.)$. If the family is reduced to a single member g one says that g is sinvex. When one can take $A = 0$ one says that g is invex.

PROPOSITION 3.4. *Let $\overline{x} \in X$. Suppose the family $\{g, -h\}$ is simultaneously sinvex. Then $\overline{x}$ is a global minimizer of f.*

The proof is obvious. Let us note that the Taylor formula ensures that both conditions are satisfied if h is quadratic and g is quadratic or convex with $h'(\overline{x}) \in \partial g(\overline{x})$, $h''(\overline{x}) \in \partial^2 g(\overline{x}, h'(\overline{x}))$. It suffices to take $v(x) = x - \overline{x}, v^*(x) = h'(\overline{x}), A(x) = h''(\overline{x})$.

4. Conclusion

Our study has taken the decomposition (1) into account as much as possible. Let us note that this decomposition suggests to introduce a notion of pseudo-critical point: $\overline{x}$ is said to be a pseudo-critical point of f if

$$0 \in \partial g(\overline{x}) \overset{*}{-} \partial h(\overline{x}).$$

This notion is weaker than the notion of critical point requiring $0 \in \partial f(\overline{x})$. However, when X is a Banach space, h is convex lower semicontinuous or if $h'(\overline{x}, \cdot) = h^*(\overline{x}, \cdot)$ and h is tangentially convex, both notions coincide, as is easily seen.

In general the inclusion

$$\partial f(\overline{x}) \subset \partial g(\overline{x}) \overset{*}{-} \partial h(\overline{x})$$

is strict, as the following example shows.

Let $X = \mathbb{R}$ and let f, g, h be given by $f(x) = |x| \sin^2 x$, $g(x) = |x|$, $h(x) = |x| \cos^2 x$. Then $\partial f(0) = \partial h(0) = \{0\}$, $\partial g(0) = [-1, 1]$ and $\partial f(0) \neq \partial g(0) \overset{*}{-} \partial h(0)$.

Since g and h may have a special structure (for instance be convex) it may be easier to deal with such a notion than with the genuine notion of critical point. We are thus lead to the following Palais-Smale condition:

($\mathrm{PS}_{d.c.}$) Any sequence (x_n) such that $(f(x_n))$ converges and there exists $x_n^* \in \partial g(x_n) \overset{*}{-} \partial h(x_n)$ with $(\|x_n^*\|) \to 0$ has a converging subsequence.

THEOREM 4.1. *Suppose X is an Asplund space, $f = g - h$ is lower semicontinuous, bounded below and satisfies condition ($\mathrm{PS}_{d.c.}$). Then f is coercive and attains its infimum.*

Proof. It suffices to observe that Condition ($\mathrm{PS}_{d.c.}$) implies the Palais-Smale condition of [17] and to apply the results of that paper. □

References

1. Asplund, E. (1973), Differentiability of the Metric Projection in Finite-Dimensional Euclidean Space, *Proceedings of the American Mathematical Society* 38, 218–219.
2. Ben Tal, A. and Zowe, J. (1982), A Unified Theory of First and Second Order Conditions for Extremum Problems in Topological Vector Spaces, *Mathematical Programming Study* 19, 39–76.
3. Bougeard, M. and Penot, J.-P. (1988), Approximation and Decomposition Properties of Some Classes of Locally d.c. Functions, *Mathematical Programming* 41, 195–227.

4. Cominetti, R. (1991), On Pseudo-differentiability, *Transactions of the American Mathematical Society* 324, 843–865.
5. Fabian, M. (1989), Subdifferentiability and Trustworthiness in the Light of a New Variational Principle of Borwein and Preiss, *Acta Universitatis Carolinae* 30, 51–56.
6. Hiriart-Urruty, J.-B. (1985), Generalized Differentiability, Duality and Optimization for Problems Dealing with Differences of Convex Functions, *Lecture Notes in Economics and Mathematical Systems 256*, Springer Verlag, Berlin, pp. 37–69.
7. Hiriart-Urruty, J.-B. (1989), Conditions nécessaires et suffisantes d'optimalité globale en optimisation de différences de fonctions convexes, *C.R. Académie des Sciences Paris* 309 I, 454–462.
8. Ioffe, A.D. (1983), On Subdifferentiability Spaces, *Annals of the New York Academy of Science* 410, 107–119.
9. Michel, Ph. (1974), Problème d'optimisation défini par des fonctions qui sont sommes de fonctions convexes et de fonctions dérivables, *Journal de Mathématiques Pures et Appliquées* 53, 321–330.
10. Moreau, J.-J. (1970), Inf-convolution, sous-additivité, convexité des fonctions numériques, *Journal de Mathématiques Pures et Appliquées* 49, 109–154.
11. Mordukhovich, B. and Shao, Y. (1996), Nonsmooth Sequential Analysis in Aplund Spaces, *Transactions of the American Mathematical Society* 348, 1235–1280.
12. Penot, J.-P. (1984), Generalized Higher Order Derivatives and Higher Order Optimality Conditions, preprint, University Santiago.
13. Penot, J.-P. (1992), Second Order Generalized Derivatives: Comparison Between Two Types of Epi-derivatives, in *Advances in Optimization, Proc. Lambrecht, FRG, 1991*, W. Oettli, D. Pallaschke, (eds.), *Lecture Notes in Economics and Mathematical Systems 382,* Springer Verlag, Berlin, pp. 52–76.
14. Penot, J.-P. (1994), Sub-hessians, Super-hessians and Conjugation, *Non-linear Analysis, Theory, Methods and Applications* 23, 689–702.
15. Penot, J.-P. (1996), Favorable Classes of Mappings and Multimappings in Nonlinear Analysis and Optimization, *Journal of Convex Analysis* 3(1), 97–116.
16. Penot, J.-P. (1993), Yet Another Mean Value Theorem, preprint, University of Pau, to appear in *J. Optimization Theory and Applications* (1997).
17. Penot, J.-P. (1994), Palais-Smale Condition and Coercivity, preprint, University of Pau.
18. Plazanet, Ph. Optimization et régularisation des différences de fonctions convexes, Thesis, University of Toulouse.
19. Rockafellar, R.T. (1982), Favorable Classes of Lipschitz Continuous Functions in Subgradient Optimization, in E. Nurminski (ed.), *Progress in Nondifferentiable Optimization*, IIASA, Laxenburg, Austria, pp. 125–144.
20. Rockafellar, R.T. (1988), First and Second-order Epi-differentiability in Nonlinear Programming, *Transactions of the American Mathematical Society* 307, 75–108.
21. Rockafellar, R.T. (1989), Second-order Optimality Conditions in Nonlinear Programming Obtained by Way of Epi-derivatives, *Mathematics of Operations Research* 14, 460–484.
22. Szulkin, A. (1986), Minimax Principles for Lower Semicontinuous Functions and Applications to Nonlinear Boundary Value Problems, *Annales Institut H. Poincaré, Analyse Non Linéaire* 3, 77–109.
23. Tao, P.D. and El Bernoussi, S. (1988), Duality in d.c. (Difference of Convex Functions) Optimization. Subgradient Methods, in K.H. Hoffmann et al. (eds.), *Trends in Mathematical Optimization*, Volume 84 of International Series of Numerisches Mathematics, Birkhauser, Basel, pp. 277–293.
24. Tao, P.D. and El Bernoussi, S. (1989), Numerical Methods for Solving a Class of Global Nonconvex Optimization Problems, in J.-P. Penot (ed.), *New Methods in Optimization and Their Industrial Uses*, Volume 87 of International Series of Numerical Mathematics, Birkhauser, Basel, pp. 97–132.
25. Tao, P.D. and An, L.T.H. (1995), Lagrangian Stability and Global Optimality in Nonconvex Quadratic Minimization over Euclidean Balls and Spheres, *Journal of Convex Analysis* 2, 263–276.
26. Thach, P.T. (1993), d.c. Sets, d.c. Functions and Nonlinear Equations, *Mathematical Programming* 58, 415–428.

27. Toland, J. (1978), Duality in Nonconvex Optimization, *Journal of Mathematical Analysis and Applications* 66, 399–415.
28. Toland, J. (1979), On Subdifferential Calculus and Duality in Nonconvex Optimization, *Bulletin Société Mathématique de France, Mémoire* 60, 173–180.
29. Tuy, H. (1987), Global Minimization of a Difference of Two Convex Functions, *Mathematical Programming Study* 30, 150–182.