Applied Mathematics E-Notes, 10(2010), 52-60 © ISSN 1607-2510
Available free at mirror sites of http://www.math.nthu.edu.tw/~amen/

Convex Functionals With Good Asymptotic Behavior On A Subset*

Mounir El Maghri†, Bouchaïb Radi‡

Received 28 November 2008

Abstract

We characterize the class of convex functionals that are asymptotically well behaved on a given convex set, instead of the whole space, by use of restriction functions. For subsets of (infinite) inequality system, the sequential stationarity becomes a Kuhn-Tucker system partially approximated. However, relaxing the condition of complementarity, we show that it does not change the class. Application to a new interior penalty method, penalizing the infinite inequalities, is given to illustrate the relevance of this result.

1 Introduction

The stationary sequences for convex functionals are reconsidered and investigated in [3] to remedy the question of unbounded sequences, those precisely having no cluster points when they are generated by a certain algorithm conceived to solve problems of optimization. Convexity is not sufficient to ensure that such sequences are minimizing. In fact, the quality of good asymptotic behavior is needed. There has been characterized a large class of closed convex functions having the property that every stationary sequence is minimizing, the so-called asymptotically well behaved functions (see [2, 3, 4]). Besides the inf-compact functions, the class contains functions that are not inf-compact. This constitutes an extension of the Palais-Smale condition in a sense that no limit points of the stationary sequences are required. The characteristic condition is obtained with the help of sublevel sections strictly above the infimum. Then, dual characterizations more easily checkable were obtained (e.g., [2, 6]). In [14, 15] are established some important links with the well-known concepts of conditioning and well-posedness. Several applications enlarging the scope of convergence for certain classical numerical methods in unbounded cases can also be found (e.g., [1], [2] and [3]). In [2], the convergence of the quadratic exterior penalty method was obtained via new Fenchel duality results for a subclass where, by definition, the stationary sequences not only are minimizing but also converge towards the optimal set assumed to be nonempty.

*Mathematics Subject Classifications: 90C25, 90C48, 90C59, 93C25.

†Department of Mathematics and Computer, Faculty of Sciences, Chouaïb Doukkali University, BP. 20, El Jadida, Morocco. E-mail: elmaghri@yahoo.com

‡Department of Mathematics, Faculty of Sciences and Technics, Hassan Premier University, Settat, Morocco. E-mail: bouchaib.radi@yahoo.fr

The purpose of this paper is that the good asymptotic behavior concept may also be taken over a convex subset instead of the whole space, in a sense that, every stationary sequence of a restricted convex function is minimizing over the restriction set. In instance of mathematical programming sets, the sequential stationarity while made explicit, leads to a KT system only partially approximated. The first result is that, we can weaken asymptotically the complementarity condition without modifying the good behavior at the infinity. The class thus remains intact. Then we introduce a new version of the logarithmic interior penalty method, actually penalizing infinitely many inequalities, and prove its convergence for this broad class, so that, the produced sequences even having no limit point are minimizing over such constraints.

2 Definitions and Preliminaries

Let X be a reflexive Banach space. For a convex function $f : X \to \mathbb{R} \cup \{+\infty\}$, let us recall the following sets of convex analysis:

- $\operatorname{dom} f = \{x \in X : f(x) \in \mathbb{R}\}$ its effective domain,
- $X_\lambda(f) = \{x \in X : f(x) \leq \lambda\}$ its sublevel set at height $\lambda \in \mathbb{R}$,
- $\partial f(\bar{x}) = \{c \in X^* : f(x) \geq f(\bar{x}) + \langle c, x - \bar{x}\rangle, \forall x \in X\}$ its subdifferential at a point $\bar{x} \in \operatorname{dom} f$ (by convention $\partial f(\bar{x}) = \emptyset$ for $\bar{x} \notin \operatorname{dom} f$), where $\langle \cdot, \cdot \rangle$ stands for the duality scalar product between X and its topological dual X^*,
- $\Gamma_0(X) = \{f : X \to \mathbb{R} \cup \{+\infty\}$ that are convex, proper and l.s.c$\}$.

A sequence $(x^n) \subset X$ is said to be stationary (for f) if $d(0, \partial f(x^n)) \to 0$. Following [3], by definition $f \in \Gamma_0(X)$ has a good asymptotic behavior (on X) if every stationary sequence is minimizing, i.e., for $(x^n) \subset X$,

$$d(0, \partial f(x^n)) \to 0 \Rightarrow f(x^n) \to \inf_X f(x).$$

This broad class of functions, denoted by $\mathcal{F}$, includes functions that are not inf-compact and do not attain their infima eventually, and has the following characterization.

THEOREM 1 ([2] or [3]).

$$\begin{aligned} \mathcal{F} &= \{f \in \Gamma_0(X) : r_\lambda(f) > 0, \forall \lambda > \inf_X f(x)\} \\ &= \{f \in \Gamma_0(X) : l_\lambda(f) > 0, \forall \lambda > \inf_X f(x)\} \end{aligned}$$

where the parameters $r_\lambda(f)$ and $l_\lambda(f)$ are defined for each scalar $\lambda > \inf_X f(x)$ by:

$$r_\lambda(f) = \inf_{f(x)=\lambda} d(0, \partial f(x)), \qquad l_\lambda(f) = \inf_{f(x)>\lambda} \frac{f(x) - \lambda}{d(x, X_\lambda(f))}.$$

It is quite natural to define the previous concept on a subset as follows. Denote first by $f_{|S}$ the restriction of $f : X \to \mathbb{R} \cup \{+\infty\}$ over the subset $S \subseteq X$, defined for all $x \in X$ by:

$$f_{|S}(x) = \begin{cases} f(x) & \text{if } x \in S \cap \operatorname{dom} f \\ +\infty & \text{otherwise} \end{cases}$$

DEFINITION 1. $f \in \Gamma_0(X)$ has a good asymptotic behavior on a nonempty closed convex subset $S \subseteq X$ if $f_{|S} \in \mathcal{F}$, i.e., $S \cap \operatorname{dom} f \neq \emptyset$ and for $(x^n) \subset X$,

$$d(0, \partial f_{|S}(x^n)) \to 0 \Rightarrow f(x^n) \to \inf_S f(x).$$

Throughout the paper we shall denote by $\mathcal{F}_S$ the class of such functions. Then $\mathcal{F}_X = \mathcal{F}$.

REMARK 1. Let us note firstly that

1. $f_{|S} = f + \delta_S$ where $\delta_S : X \to \mathbb{R} \cup \{+\infty\}$ is the indicator function of S. Obviously, with S nonempty convex closed and $f \in \Gamma_0(X)$, $f_{|S} \in \Gamma_0(X)$ iff $S \cap \operatorname{dom} f \neq \emptyset$.
2. $d(0, \partial f_{|S}(x^n)) \to 0$ means that $\forall n$, $\exists c^n \in \partial f_{|S}(x^n) : \|c^n\| \to 0$. In this case, we must have $x^n \in S \cap \operatorname{dom} f$ $(\forall n)$ otherwise $\partial f_{|S}(x^n) = \emptyset$. So $f_{|S}(x^n) = f(x^n)$.

So it suffices to apply Theorem 1 to the function $f + \delta_S$ (using Remark 1) to obtain the characterization of the class $\mathcal{F}_S$ in the general case of implicit subsets.

COROLLARY 1.

$$\begin{aligned} \mathcal{F}_S &= \{f \in \Gamma_0(X) \,|\, S \cap \operatorname{dom} f \neq \emptyset : r_\lambda(f_{|S}) > 0, \forall \lambda > \inf_S f(x)\} \\ &= \{f \in \Gamma_0(X) \,|\, S \cap \operatorname{dom} f \neq \emptyset : l_\lambda(f_{|S}) > 0, \forall \lambda > \inf_S f(x)\} \end{aligned}$$

where the parameters $r_\lambda(f_{|S})$ and $l_\lambda(f_{|S})$ are given for each real $\lambda > \inf_S f(x)$ by:

$$r_\lambda(f_{|S}) = \inf_{\substack{f(x)=\lambda \\ x \in S}} d(0, \partial(f + \delta_S)(x)), \qquad l_\lambda(f_{|S}) = \inf_{\substack{f(x)>\lambda \\ x \in S}} \frac{f(x) - \lambda}{d(x, S_\lambda(f))}$$

and $S_\lambda(f) = \{x \in S : f(x) \leq \lambda\}$ is as usual the λ-sublevel set of f restricted to S.

3 The Explicit Case of Mathematical Programming

We henceforth assume that the set S is of mathematical programming form:

$$S = \{x \in X : G(x) \in -Y_+\} \tag{1}$$

where $G : X \to Y \cup \{+\infty\}$ taking values in Y a Banach space equipped with a partial order induced by a closed convex cone $Y_+ \subset Y$: $\forall y, y' \in Y$,

$$y \leq_{Y_+} y' \;\Leftrightarrow\; y - y' \in -Y_+.$$

Then S can be rewritten with $G(x) \leq_{Y_+} 0$. The element $+\infty$ is adjoined to Y to be its greatest element: $\forall y \in Y$, $y \leq_{Y_+} +\infty$. The effective domain of the vector mapping G is defined by $\operatorname{dom} G = \{x \in X : G(x) \in Y\}$. The vector mapping G is said to be

- Y_+-convex, if

  $\forall x, x' \in X$, $\forall \alpha \in [0,1]$, $G(\alpha x + (1-\alpha)x') \leq_{Y_+} \alpha G(x) + (1-\alpha)G(x')$,
- sequentially Y_+-l.s.c at $\bar{x} \in X$, if

  $\forall y \leq_{Y_+} G(\bar{x})$, $\forall (x^n) \to \bar{x}$, $\exists (y^n) \to y :\; y^n \leq_{Y_+} G(x^n)$, $\forall n \in \mathbb{N}$.

Sequential Y_+-l.s.c (at every $\bar{x} \in X$) easily implies that the epigraph and sublevel sets are closed; the converse fails [7]. In particular if G is sequentially Y_+-l.s.c then S is closed. In a metrizable space, in particular the Banach space X, sequential $\mathbb{R}_+$-l.s.c is no more than the classical l.s.c, and then by definition, sequential $\mathbb{R}_+^m$-l.s.c also becomes equivalent to l.s.c of the components. We shall adopt for such a vector mapping a similar notation to a scalar function:

$$\Gamma_0(X,Y) = \{G : X \to Y \cup \{+\infty\}\ Y_+\text{-convex proper sequentially } Y_+\text{-l.s.c}\}.$$

A composite function $\psi \circ \varphi : X \to \mathbb{R} \cup \{+\infty\}$ is defined by $(\psi \circ \varphi)(x) = \psi(\varphi(x))$ if $x \in \operatorname{dom} \varphi$ and $+\infty$ otherwise, and its domain is therefore

$$\operatorname{dom}(\psi \circ \varphi) = \operatorname{dom} \varphi \cap \varphi^{-1}(\operatorname{dom} \psi). \tag{2}$$

The following formula was established in [7, pp.135]: Let $f \in \Gamma_0(X)$, $G \in \Gamma_0(X,Y)$ be such that the following qualification condition of Attouch-Brézis type is satisfied:

$$(AB) \qquad \mathbb{R}_+[Y_+ + G(\operatorname{dom} f \cap \operatorname{dom} G)] \quad \text{is a closed vector subspace of } Y.$$

Then for all $\bar{x} \in S \cap \operatorname{dom} f$,

$$\partial f_{|S}(\bar{x}) = \bigcup_{\substack{\mu \in Y_+^* \\ \langle \mu, G(\bar{x}) \rangle = 0}} \partial \mathcal{L}(\bar{x}, \mu) \tag{3}$$

where $Y_+^* = \{\mu \in Y^* : \langle \mu, y \rangle \geq 0, \forall y \in Y_+\}$ is the (positive) polar cone of Y_+. In fact in [7], $f_{|S} = f + \delta_{-Y_+} \circ G$ and $\mathcal{L}(., \mu) = f + \mu \circ G$ the well-known Lagrangian.

REMARK 2. The above formula well known before [13], actually requires a constraint qualification weaker than the Slater condition "$\exists a \in \operatorname{dom} f,\ G(a) \in -\operatorname{int} Y_+$". Indeed, the latter easily implies that $\mathbb{R}_+[Y_+ + G(\operatorname{dom} f \cap \operatorname{dom} G)] = Y$. Furthermore, the condition (AB) ensures that $0 \in C = Y_+ + G(\operatorname{dom} f \cap \operatorname{dom} G)$, that is, $S \cap \operatorname{dom} f \neq \emptyset$. Indeed, observe that for any nonempty convex set C, the set $\mathbb{R}_+ C$ is a vector space if and only if $]0, +\infty[C$ is a vector space. Thus $0 \in C$.

It is now easy to verify that for $f \in \Gamma_0(X)$ and $G \in \Gamma_0(X,Y)$ under (AB) condition, $(x^n) \subset X$ is a stationary sequence for $f_{|S}$ iff it is of Kuhn-Tucker type:

$$(KT) \quad \begin{cases} \text{there exists a sequence } (\mu^n) \subset Y_+^* \text{ of Lagrange multipliers such that,} \\ d(0, \partial \mathcal{L}(x^n, \mu^n)) \to 0 \qquad \text{(asymptotic Lagrangian stationarity)} \\ \langle \mu^n, G(x^n) \rangle = 0 \quad (\forall n) \qquad \text{(complementarity)} \\ G(x^n) \in -Y_+ \quad (\forall n) \qquad \text{(feasibility)} \end{cases}$$

Let us relax asymptotically the complementarity condition:

$$(\widetilde{KT}) \quad \begin{cases} \text{there exists a sequence } (\mu^n) \subset Y_+^* \text{ of Lagrange multipliers such that,} \\ d(0, \partial \mathcal{L}(x^n, \mu^n)) \to 0 \qquad \text{(asymptotic Lagrangian stationarity)} \\ \langle \mu^n, G(x^n) \rangle \to 0 \qquad \text{(asymptotic complementarity)} \\ G(x^n) \in -Y_+ \quad (\forall n) \qquad \text{(feasibility)} \end{cases}$$

The property that every bounded stationary sequence is minimizing, may fail with ones unbounded, as shown by the counter-example below.

EXAMPLE 1. It has been shown in [3] that the function $f \in \Gamma_0(\mathbb{R}^2)$ defined by $f(x_1, x_2) = \frac{x_1^2}{x_2}$ if $x_2 > 0$, $= 0$ if $x_1 = x_2 = 0$, $+\infty$ elsewhere, has a bad asymptotic behavior on $\mathbb{R}^2$. We show that it still has this behavior on $S = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq 1\}$. Indeed, $x^n = (n, n^2)$ is not minimizing for f over S because $f(x^n) = 1 \not\to 0 = \inf_S f(x)$. But it is $(\widetilde{KT})$ (resp. (KT)) stationary taking $\mu^n = \frac{1}{n^3}$ (resp. $\mu^n = 0$):

$$\begin{cases} f'(x^n) + \mu^n G'(x^n) = (\frac{2}{n}, \frac{-1}{n^2} + \frac{-1}{n^3}) \to (0,0), \\ \mu^n G(x^n) = \frac{1}{n^3}(1 - n^2) \to 0. \end{cases}$$

Given S defined by (1) with $G \in \Gamma_0(X, Y)$, we consider the following two classes:

$$\mathcal{F}_S^{AB} = \{ f \in \Gamma_0(X) \text{ satisfying } (AB) : \text{every } (KT) \text{ sequence is minimizing for } f_{|S} \},$$

$$\widetilde{\mathcal{F}}_S^{AB} = \{ f \in \Gamma_0(X) \text{ satisfying } (AB) : \text{every } (\widetilde{KT}) \text{ sequence is minimizing for } f_{|S} \}.$$

Then it is immediate that

$$\widetilde{\mathcal{F}}_S^{AB} \subseteq \mathcal{F}_S^{AB} = \mathcal{F}_S \cap \{ f \in \Gamma_0(X) \text{ satisfying } (AB) \}. \tag{4}$$

As mentioned before, the class $\mathcal{F}_S^{AB}$ is in fact unchanged after the relaxation:

THEOREM 2.

$$\widetilde{\mathcal{F}}_S^{AB} = \mathcal{F}_S^{AB}.$$

PROOF. According to (4), it remains to show the converse inclusion. Let $f \in \mathcal{F}_S^{AB}$ and suppose $f \notin \widetilde{\mathcal{F}}_S^{AB}$. Then we may find (x^n) a $(\widetilde{KT})$ stationary sequence not minimizing $f_{|S}$. Hence we may find $\mu^n \in Y_+^*$ and $c^n \in \partial\mathcal{L}(x^n, \mu^n)$ for each n, $\varrho > \inf_S f(x)$ and $\lambda \in f(S)$ such that:

$$\begin{cases} \|c^n\| \to 0, \\ \langle \mu^n, G(x^n) \rangle \to 0, \\ G(x^n) \in -Y_+ \qquad \text{for all } n, \\ \inf_S f(x) < \lambda < \varrho \leq f(x^n) \qquad \text{for infinitely many } n. \end{cases}$$

Since $f \in \mathcal{F}_S^{AB}$, by (4) and Corollary 1, for such λ, we have $l_\lambda(f_{|S}) > 0$. Let p^n be a projection of x^n over $S_\lambda(f)$ (defined in Corollary 1) which is a nonempty closed convex subset of the *reflexive* Banach space X. Hence $\|x^n - p^n\| = d(x^n, S_\lambda(f))$, $f(p^n) \leq \lambda$ and $G(p^n) \in -Y_+$, and then, we have

$$\langle c^n, x^n - p^n \rangle \geq \mathcal{L}(x^n, \mu^n) - \mathcal{L}(p^n, \mu^n) \geq f(x^n) - \lambda + \langle \mu^n, G(x^n) \rangle.$$

So dividing these inequalities by $f(x^n) - \lambda$, we obtain for infinitely many n,

$$\|c^n\| \frac{1}{l_\lambda(f_{|S})} \geq \|c^n\| \frac{\|x^n - p^n\|}{f(x^n) - \lambda} \geq 1 + \langle \mu^n, G(x^n) \rangle \frac{1}{f(x^n) - \lambda}.$$

Since $(\frac{1}{f(x^n) - \lambda})$ is bounded, by letting $n \nearrow +\infty$, we get the contradiction $0 \geq 1$.

REMARK 3. The parameter $r_\lambda(f_{|S})$ of Corollary 1 can be given explicitly via (3):

$$r_\lambda(f_{|S}) = \inf_{\substack{f(x)=\lambda \\ x \in S}} \inf \{ \|c\| : c \in \bigcup_{\substack{\langle \mu, G(x) \rangle = 0 \\ \mu \in Y_+^*}} \partial\mathcal{L}(x, \mu) \} = \inf_{\substack{f(x)=\lambda \\ G(x) \in -Y_+ \\ \langle \mu, G(x) \rangle = 0 \\ \mu \in Y_+^*}} d(0, \partial\mathcal{L}(x, \mu)).$$

4 Interior Penalty Method

Consider the oriented distance [12] defined for the closed convex cone $-Y_+$ by:

$$\Delta_{-Y_+}(y) = d(y, -Y_+) - d(y, Y \setminus -Y_+).$$

This function is convex positively homogeneous, 1-Lipschitzian, Y_+-nondecreasing and characterizes the boundary, interior and complementary of $-Y_+$:

$$\partial(-Y_+) = \{y \in Y : \Delta_{-Y_+}(y) = 0\}, \quad \mathrm{int}(-Y_+) = \{y \in Y : \Delta_{-Y_+}(y) < 0\},$$
$$Y \setminus (-Y_+) = \{y \in Y : \Delta_{-Y_+}(y) > 0\}.$$

It is also known [11] that the function Δ_{-Y_+} is simply given for all $y \in Y$ by:

$$\Delta_{-Y_+}(y) = \sup_{\mu \in Y_+^*,\ \|\mu\|=1} \langle \mu, y \rangle. \tag{5}$$

The feasible set S defined by (1) can now be written under scalarized form:

$$S = \{x \in X : \Delta_{-Y_+}(G(x)) \leq 0\}.$$

The classical logarithmic penalty function applied with f over S becomes:

$$f_n = f - r_n \ln \circ (-\Delta_{-Y_+} \circ G), \qquad r_n \searrow 0^+.$$

Standard theorem for this penalization applies under the classical hypotheses of (inf-)compactness type. We shall extend the convergence for the class $\mathcal{F}_S^{AB}$.

The result below is needed.

THEOREM 3. [7] Let Z be a topological vector space ordered by a convex cone Z_+, $f : X \to \mathbb{R} \cup \{+\infty\}$ be convex proper, $H : X \to Z \cup \{+\infty\}$ be Z_+-convex proper, $g : Z \to \mathbb{R} \cup \{+\infty\}$ be convex proper Z_+-nondecreasing and one of the following qualification conditions (resp. of Moreau-Rockafellar and Attouch-Brézis type) hold:

$$\begin{cases} X, Z \text{ are locally convex spaces,} \\ g \text{ is finite and continuous at some point of } H(\mathrm{dom}\, f \cap \mathrm{dom}\, H), \end{cases}$$

or,

$$\begin{cases} X, Z \text{ are Fréchet spaces, } H \text{ is sequentially } Z_+\text{-l.s.c,} \\ \mathbb{R}_+[\mathrm{dom}\, g - H(\mathrm{dom}\, f \cap \mathrm{dom}\, H)] \text{ is a closed vector subspace of } Z. \end{cases}$$

Then $\forall \bar{x} \in X$,

$$\partial(f + g \circ H)(\bar{x}) = \bigcup_{\mu \in \partial g(H(\bar{x}))} \partial(f + \mu \circ H)(\bar{x}).$$

We derive the same result in concordance with our data.

REMARK 4. The above formula still holds if we assume that g is Z_+-nonincreasing and $H : X \to Z \cup \{-\infty\}$ is Z_+-concave sequentially Z_+-u.s.c. It suffices to apply Theorem 3 to the functions $g \circ -Id_Z$ (Id_Z is the identity mapping) and $-H$ (instead of g and H) observing that

$$\partial(g \circ -Id_Z)(-H(\bar{x})) = -\partial g(H(\bar{x})).$$

Recall for $\psi : Y \to \mathbb{R} \cup \{+\infty\}$ convex proper and $y \in \operatorname{dom}\psi$, the Young-Fenchel's equality:

$$y^* \in \partial\psi(y) \Longleftrightarrow \psi(y) + \psi^*(y^*) = \langle y^*, y\rangle \tag{6}$$

the function $\psi^* : Y^* \to \mathbb{R} \cup \{+\infty\}$ being the Legendre-Fenchel's conjugate of ψ defined by $\psi^*(y^*) = \sup_{y\in Y}\{\langle y^*, y\rangle - \psi(y)\}$. We have that $\psi^{**} = \psi$ if $\psi \in \Gamma_0(Y)$ (see e.g. [5]).

The symbol "co" will stand for the convex hull and B for the closed unit ball.

LEMMA 1. For the convex cone Y_+ (not necessarily closed here), we have that

1. $\Delta^*_{-Y_+} = \delta_{\overline{\mathrm{co}}(Y^*_+\cap\partial B)}$,
2. $\partial\Delta_{-Y_+}(\bar y) = \{\mu \in Y^* : \langle \mu, \bar y\rangle = \Delta_{-Y_+}(\bar y),\ \mu \in \overline{\mathrm{co}}(Y^*_+ \cap \partial B)\}$ for all $\bar y \in Y$.

PROOF. 1. By (5) (which holds for any convex cone) we have for every $y \in Y$,

$$\begin{aligned}\Delta_{-Y_+}(y) = \sup_{\mu\in Y^*_+\cap\partial B}\langle \mu, y\rangle &= \sup_{\mu\in \mathrm{co}(Y^*_+\cap\partial B)}\langle \mu, y\rangle\\ &= \sup_{\mu\in \overline{\mathrm{co}}(Y^*_+\cap\partial B)}\langle \mu, y\rangle\\ &= \sup_{\mu\in Y^*}\{\langle \mu, y\rangle - \delta_{\overline{\mathrm{co}}(Y^*_+\cap\partial B)}(\mu)\} = \delta^*_{\overline{\mathrm{co}}(Y^*_+\cap\partial B)}(y).\end{aligned}$$

The indicator function $\delta_{\overline{\mathrm{co}}(Y^*_+\cap\partial B)} \in \Gamma_0(Y^*)$ because $\overline{\mathrm{co}}(Y^*_+\cap\partial B)$ is a nonempty convex closed set in Y^*. Hence $\delta_{\overline{\mathrm{co}}(Y^*_+\cap\partial B)} = \delta^{**}_{\overline{\mathrm{co}}(Y^*_+\cap\partial B)} = \Delta^*_{-Y_+}$.

2. This assertion follows directly from (6) and the fact that $\operatorname{dom}\Delta_{-Y_+} = Y$.

REMARK 5. As first consequences, we have

1. $\partial\Delta_{-Y_+}(\bar y) \subset Y^*_+ \setminus \{0\} \cap B$ for all $\bar y \in Y$, if $\operatorname{int} Y_+ \neq \emptyset$.

   Indeed, $Y^*_+ \cap \partial B \subset Y^*_+ \cap B$. Now, if $0 \in \partial\Delta_{-Y_+}(\bar y)$ then we get $\Delta_{-Y_+}(y) \geq 0$ ($\forall y \in Y$) contradicting the fact that $\Delta_{-Y_+}(y) < 0$ ($\forall y \in -\operatorname{int} Y_+$).
2. $\partial\Delta_{-Y_+}(\bar y) = \{\mu \in Y^* : \mu \in Y^*_+,\ \|\mu\| = 1,\ \langle \mu, \bar y\rangle = \Delta_{-Y_+}(\bar y)\}$ for all $\bar y \notin -Y_+$.

   Indeed, according to the first remark, it suffices to show that $\|\mu\| \geq 1$. But this follows easily from the fact that $0 < \Delta_{-Y_+}(\bar y) = \langle \mu, \bar y\rangle \leq \|\mu\|\Delta_{-Y_+}(\bar y)$.

PROPOSITION 1. With f proper convex and G proper Y_+-convex, we have, $\forall\, n$, $\forall\, \bar x \in \operatorname{dom} f \cap G^{-1}(-\operatorname{int} Y_+)$,

$$\partial f_n(\bar x) = \bigcup_{\substack{\mu\in\overline{\mathrm{co}}(Y^*_+\cap\partial B)\\ \langle\mu, G(\bar x)\rangle = \Delta_{-Y_+}(G(\bar x))}} \partial\Big(f - \frac{r_n}{\Delta_{-Y_+}(G(\bar x))}\mu \circ G\Big)(\bar x). \tag{7}$$

PROOF. If $\bar x \notin \operatorname{dom} f_n$ then by definition $\partial f_n(\bar x) = \emptyset$. By (2) we have that

$$\bar x \in \operatorname{dom} f_n = \operatorname{dom} f \cap \{x \in X : G(x) \in -\operatorname{int} Y_+\}. \tag{8}$$

According to Remark 4, the functionals f, $g = -r_n \ln$ and $H = -\Delta_{-Y_+} \circ G$ satisfy the hypotheses of Theorem 3 with the Moreau-Rockafellar condition. Hence

$$\partial f_n(\bar x) = \partial\Big(f - \frac{r_n}{\Delta_{-Y_+}(G(\bar x))}\Delta_{-Y_+} \circ G\Big)(\bar x).$$

The functions f, $g = -\frac{r_n}{\Delta_{-Y_+}(G(\bar{x}))}\Delta_{-Y_+}$ and $H = G$ also satisfy the hypotheses of Theorem 3 with the Moreau-Rockafellar type condition. Hence

$$\partial f_n(\bar{x}) = \bigcup_{\mu \in -\frac{r_n}{\Delta_{-Y_+}(G(\bar{x}))}\partial \Delta_{-Y_+}(G(\bar{x}))} \partial(f + \mu \circ G)(\bar{x}).$$

Finally by using Lemma 1, we obtain the formula of the proposition.

The convergence of this penalty method for the class $\mathcal{F}_S^{AB}$ can now be announced as follows.

THEOREM 4. Let $f \in \mathcal{F}_S^{AB}$. Then, every diagonally stationary sequence for (f_n) is minimizing for $f_{|S}$, i.e.,

$$d(0, \partial f_n(x^n)) \to 0 \Rightarrow f(x^n) \to \inf_S f(x).$$

PROOF. The proof consists in showing that (x^n) is a $(\widetilde{KT})$ sequence for $f_{|S}$. Indeed, as in 2^o) Remark 1, $x^n \in \operatorname{dom} f_n$ for all n. So by (7, 8) and 1^o) Remark 5, we can deduce the existence of a sequence of Lagrange multipliers $(\mu^n) \subset Y_+^* \setminus \{0\}$ such that,

$$\begin{cases} d(0, \partial \mathcal{L}(x^n, -\frac{r_n}{\Delta_{-Y_+}(G(x^n))}\mu^n)) \to 0, \\ \langle \mu^n, G(x^n) \rangle = \Delta_{-Y_+}(G(x^n)) \quad (\forall n), \\ G(x^n) \in -\operatorname{int} Y_+ \quad (\forall n). \end{cases}$$

Renaming $-\frac{r_n}{\Delta_{-Y_+}(G(x^n))}\mu^n$ by μ^n which still lies in the cone $Y_+^* \setminus \{0\}$, and using the fact that $(r_n) \searrow 0^+$, we obtain the asymptotic complementarity condition:

$$\langle \mu^n, G(x^n) \rangle = -r_n \nearrow 0^-.$$

So (x^n) is a $(\widetilde{KT})$ sequence. It follows by Theorem 2 that (x^n) is minimizing for $f_{|S}$.

Acknowledgments. The authors would like to thank the reviewer very much for his useful suggestion and his considerable comments.

References

[1] A. Auslender, Asymptotic properties of the Fenchel dual functional and applications to decomposition problems, J. Optim. Theory Appl., 73(1992), 427–449.

[2] A. Auslender, R. Cominetti and J. P. Crouzeix, Convex functions with unbounded level sets and applications to duality theory, SIAM J. Optim., 3(1993), 669–687.

[3] A. Auslender and J. P. Crouzeix, Well behaved asymptotical convex functions, Ann. Inst. Poincaré. Anal. Non Linéaire, 6(1989), 101–122.

[4] A. Auslender and M. Teboulle, Asymptotic Cones and Functions in Optimization and Variational Inequalities, Springer-Verlag, New York, 2003.

[5] D. Azé, Eléments d'Analyse Convexe et Variationnelle, Ellipses, Paris, 1997.

[6] D. Azé and L. Michel, Computable dual characterizations of asymptotically well-behaved convex functions, Comm. Appl. Nonlinear Analysis, 6(1999), 39–49.

[7] C. Combari, M. Laghdir and L. Thibault, Sous-différentiels de fonctions convexes composées, Ann. Sci. Math. Québec, 18(1994), 119–148.

[8] A. Coulibaly and J. P. Crouzeix, Condition numbers and error bounds in convex programming, Math. Program., Ser. B, 116(2009), 79–113.

[9] M. El Maghri and B. Bernoussi, Pareto optimizing and Kuhn–Tucker stationary sequences, Numer. Funct. Anal. Optim., 28(2007), 287–305.

[10] M. El Maghri and M. Laghdir, Pareto subdifferential calculus for convex vector mappings and applications to vector optimization, SIAM J. Optim. 19(4)(2009), 1970–1994.

[11] I. Ginchev and A. Hoffmann, Approximation of set-valued functions by single-valued one, Discuss. Math., Diff. Inclus., Control and Optim., 22(2002), 33–66.

[12] J. B. Hiriart-Urruty, Tangent cones, generalized gradients and mathematical programming in Banach spaces, Math. Oper. Res., 4(1979), 79–97.

[13] B. Lemaire, Application of a subdifferential of a convex composite functional to optimal control in variational inequalities, Lect. Notes Eco. Math. Syst., Springer-Verlag, New York, 255(1985), 103–117.

[14] B. Lemaire, Bonne position, conditionnement et bon comportement asymptotique, Séminaire d'Analyse Convexe, Exposé n^{0}5, Montpellier, France, 1992.

[15] J. P. Penot, Well posedness and nonsmooth analysis, Pliska Stus. Math. Bulgar., (12)1998, 141–190.