Journal of Global Optimization **10:** 305–326, 1997.

Hunting for a Smaller Convex Subdifferential

V. F. DEMYANOV* and V. JEYAKUMAR
Applied Mathematics Department, St. Petersburg State University, Staryi Peterhof 198904, Russia; School of Mathematics, University of New South Wales, Sydney 2052, Australia.

(Received: 28 February 1996; accepted: 2 July 1996)

Abstract. Certain useful basic results of the gradient (in the smooth case), the Clarke subdifferential, the Michel–Penot subdifferential, which is also known as the "small" subdifferential, and the directional derivative (in the nonsmooth case) are stated and discussed. One of the advantages of the Michel–Penot subdifferential is the fact that it is in general "smaller" than the Clarke subdifferential. In this paper it is shown that there exist subdifferentials which may be smaller than the Michel–Penot subdifferential and which have certain useful calculus. It is further shown that in the case of quasidifferentiability, the Michel–Penot subdifferential enjoys calculus which hold for the Clarke subdifferential only in the regular case.

Key words: Clarke subdifferential, Michel–Penot subdifferential, quasidifferential, convexificator, mean-value theorem, nonsmooth analysis.

AMS Subject Classification: 90C30, 90C48, 49B27.

1. Introduction

With the discovery of the convex subdifferential (see [16]) and the subdifferential of a max-type function (see, e.g. [3,4]) it was generally understood that in the nonsmooth case it is not sufficient to employ a singleton – the gradient – to study properties of a function. Since the subdifferentials (in the mentioned cases) appeared to be convex sets it was a great temptation to look for similar convex ones in a general nonsmooth case. Pschenichnyi [13] developed upper convex and lower concave approximations. The introduction of the Clarke sub-differential [1] was a great breakthrough, and a safari season started in the Wilderness of Endolandia.[1] Many different generalizations of the concept of gradient have been proposed. The most productive hunter (as to the authors' knowledge) is J.-P. Penot. He discovered and studied many convex objects, one of the most promising and popular being that of "small subdifferential" (nurtured jointly by P. Michel and J.-P. Penot [12]).

Our aim is two-fold:
(a) to demonstrate that in some cases there exists a Calculus for "small" subdifferentials; (b) to show that, like a kangaroo, the small subdifferential contains in itself even a smaller one (or smaller ones).

* Research of the first author was partially supported by grants from the Australian Research Council and the G. Soros Foundation (p.468)

[1] The Land of NDO-Nondifferentiable Optimization – the term introduced by M. Balinski

Indeed, the Wilderness of Endolandia is still full of surprises and a persistant hunter may be lucky.

In the paper only the finite-dimenstional case is considered. We discuss mostly Lipschitz directionally differentiable functions. After a brief review of the properties of the gradient in the smooth case we show that in the nonsmooth case several tools are needed to solve the problems which in the smooth case are solved by means of the gradient. In particular, we indicate several problems where the Clarke subdifferential can be employed for solving them. Then we observe that the same problems can be treated by means of the Michel–Penot subdifferential which in some cases is "smaller" than the Clarke one and, hence, provides sharper results. It is also shown that for the Michel–Penot subdifferentials there exists (in the case of Lipschitz quasidifferentiable functions) a Calculus (exact rules – equalities – for computing "small" subdifferentials).

Using the idea of convexificator [7,8] we can solve the same problems which are solved by means of the Clarke and Michel–Penot subdifferentials. It turns out that both subdifferentials (the Clarke one and Michel–Penot one) are convexifactors and have been used just in this capacity. In some cases it is possible to find a convexificator (or convexificators) which is (are) even smaller than the Michel–Penot one. For this purpose the notion of minimal convexificator seems to be promising [7,8]. It opend the way for a more thorough study of nonsmooth functions.

2. The Gradient: The Principal Tool in Smooth Analysis

We begin by recalling some properties of the gradient – the principal tool in smooth analysis.

Let a function $f : \Omega \to \mathbb{R}$ be continuously differentiable on Ω where $\Omega \subseteq \mathbb{R}^n$ is an open set. Then the gradient mapping $f' : \Omega \to \mathbb{R}^n$ is defined and continuous on Ω. Fix any $x \in \Omega$. By means of the gradient one is able:

1. To find the directional derivative of f at x in any direction g :

$$f'(x,g) := \lim_{\alpha\downarrow 0} \frac{f(x+dg)-f(x)}{\alpha} = (f'(x),g). \tag{1}$$

2. To construct a first-order approximation of f near x:

$$f(x+\Delta) = f(x) + (f'(x),\Delta) + 0_x(\Delta) \tag{2}$$

where for each Δ,

$$\frac{0_x(\alpha\Delta)}{\alpha} \underset{\alpha\downarrow 0}{\longrightarrow} 0. \tag{3}$$

3. To formulate necessary conditions for an extremum:

3a: For a point $x^* \in \Omega$ to be a minimum point of f it is necessary that

$$f'(x^*) = 0_n \tag{4}$$

3b: For a point $x^{**} \in \Omega$ to be a maximum point of f it is necessary that

$$f'(x^{**}) = 0. \tag{5}$$

Note that the conditions (4) and (5) coincide and that a point x, satisfying (4) and (5) is called a stationary point.

4. To find directions of the steepest descent and ascent:

   4a: the direction $g_0 = -\dfrac{f'(x)}{\|f'(x)\|}$ is the steepest descent direction.

   4b: the direction $g_1 = \dfrac{f'(x)}{\|f'(x)\|}$ is the steepest ascent direction.

   Note that in the smooth case there exists only one steepest descent direction and only one steepest ascent direction and that

$$g_0 = -g_1. \tag{6}$$

The following also hold:

5. If $f'(x, g) < 0$ then $f'(x, -g) > 0$, that is, if a function f is decreasing in some direction it is necessarily increasing in the opposite direction, and, by the way, due to (1),

$$f'(x, g) = -f'(x, -g). \tag{7}$$

6. If $f'(x, g) < 0$ then $f'(x', g) < 0$ for all x' near the point x (that is, the direction g is a robust direction of descent: if f is decreasing at x in a direction g, it is also decreasing in the same direction at all points from some neighbourhood of x). The same is true with respect to ascent directions.
7. The function $F(x, \Delta) = (f'(x), \Delta) = f'(x, \Delta)$, which is an approximation of the increment $f(x + \Delta) - f(x)$, is continuous as a function of x.
8. The following mean-value theorem is valid: If the interval $co\{x_1, x_2\} \subset \Omega$ then there exists and $\alpha \in (0, 1)$ such that

$$f(x_2) - f(x_1) = (f'(x_1 + \alpha(x_2 - x_1)), x_2 - x_1). \tag{8}$$

9. It is also understood (usually not stated explicitly) that one can study the above properties (and many others) by means only $(n + 1)$ numbers (the value of f at x and n partial derivatives constituting the gradient). Therefore it is only necessary to compute and to store the mentioned $(n + 1)$ numbers.

3. Directionally Differentiable Functions

Now let us assume that $f : \Omega \to \mathbb{R}$ is directionally differentiable (d.d.) on Ω (i.e., the limit (1) exists and is finite for every $g \in \mathbb{R}^n$) and the directional derivative is continuous as a function of direction. Since $f'(x, g)$ is positively homogeneous (p.h.) of degree 1, i.e.

$$f'(x, \lambda g) = \lambda f'(x, g), \quad \forall \lambda \geq 0, \tag{9}$$

it is sufficient to consider only $g \in S_1 = \{g \in \mathbb{R}^n \mid \|g\| = 1\}$.

Examining problems 1–8 of Section 2, we observe that the directional derivative allows us

(1) to find the directional derivative (by the definition),

(2) to construct a first-order approximation,

$$f(x+\Delta) = f(x) + f'(x,\Delta) + 0_x(\Delta) \tag{10}$$

where $0_x(\Delta)$ satisfies (3),

(3) to formulate necessary conditions for an extremum; thus,

(3a) for a point $x^* \in \Omega$ to be a minimum point of f it is necessary that

$$f'(x^*,g) \geq 0, \quad \forall g \in \mathbb{R}^n. \tag{11}$$

If

$$f'(x^*,g) > 0, \quad \forall g \neq 0_n \tag{12}$$

then x^* is a strict local minimum point.

(3b) for a point $x^{**} \in \Omega$ to be a maximum point of f it is necessary that

$$f'(x^{**},g) \leq 0 \quad \forall g \in \mathbb{R}^n. \tag{13}$$

If

$$f'(x^{**},g) < 0 \quad \forall g \neq 0_n \tag{14}$$

then x^{**} is a strict local maximum point.

Note that necessary conditions for a maximum and a minimum do **not** coincide any more and sufficient conditions (12) and (14) have no equivalence in the smooth case, they are just impossible.

(4) to define directions of steepest ascent and descent; however, they are not necessarily unique any more, and the property similar to (6) doesn't hold now. The properties (5) and (7) of Section 2 don't hold any more.

(8) to formulate the following mean-value property that for for the interval $co\{x_1,x_2\} \subset \Omega$, there exists an $\alpha \in (0,1)$ such that

$$f(x_2) - f(x_1) = f'(x_1 + \alpha(x_2 - x_1),\ x_2 - x_1). \tag{15}$$

4. Locally Lipschitz Functions

Assume now that $f : \Omega \to \mathbb{R}$ is locally Lipschitz. To preserve something similar to properties (5) - (7), one can employ upper and lower Clarke directional derivatives:

$$f_{cl}^{\uparrow}(x,g) = \limsup_{[\alpha,x'] \to [0^+,\ x]} \frac{1}{\alpha}[f(x'+\alpha g) - f(x')] \tag{16}$$

$$f_{cl}^{\downarrow}(x,g) = \liminf_{[\alpha,x']\to[0^+,\ x]} \frac{1}{\alpha}[f(x'+\alpha g) - f(x')]. \tag{17}$$

If f is locally Lipschitz then the limits in (16) and (17) exist and are finite and the following important properties hold:

$$f_{cl}^{\uparrow}(x,g) = \max_{v\in\partial_{cl}f(x)}(v,g),$$

$$f_{cl}^{\downarrow}(x,g) = \min_{w\in\partial_{cl}f(x)}(w,g), \tag{18}$$

where

$$\partial_{cl}f(x) = co\{v \in \mathbb{R}^n | \exists\{x_n\} : x_k \to x,\ x_k \in T(f),\ f'(x_k) \to v\}, \tag{19}$$

$T(f)$ is the set of points of Ω where f is differentiable (since f is locally Lipschitz, f is differentiable almost everywhere). The set $\partial_{cl}f(x)$, called the Clarke subdifferential of f at x, is a nonempty, convex and compact set (see [1]).

The inclusion $0 \in \partial_{cl}f(x)$ is a necessary optimality condition (both for a minimum and a maximum). If f is also d.d. and $0 \in \partial_{cl}f(x)$, then there exist directions $g's$ such that

$$f_{cl}^{\uparrow}(x,g) < 0. \tag{20}$$

It is clear from (18) that $f_{cl}^{\downarrow}(x,-g) = -f_{cl}^{\uparrow}(x,g) > 0$ and now it is easy to see that

$$f'(x,g) \le f_{cl}^{\uparrow}(x,g) < 0, \quad f'(x,-g) \ge f_{cl}^{\downarrow}(x,g) > 0. \tag{21}$$

Hence, the property (21) is a replacement for property (5) in the smooth case. Thus, if (21) holds, g is a descent direction and $g_1 = -g$ is an ascent direction.

It is also possible to show that if $f_{cl}^{\uparrow}(v,g) < 0$, then $f'(x,g) < 0$ and $f'(x',g) < 0$ for all x' near x, i.e. this property is similar to property (6) in the smooth case (and, hence, g is a robust direction of descent). An analogous property holds if $f_{cl}^{\downarrow}(v,g) > 0$ (then g is a robust direction of ascent).

The functions $f_{cl}^{\uparrow}(x,g)$ and $f_{cl}^{\downarrow}(x,g)$ are, in general, discontinuous (as well as the set-valued mapping $\partial_{cl}f : \mathbb{R}^n \to 2^{\mathbb{R}^n}$). A mean-value theorem can be formulated by means of the Clarke directional derivatives (see [1]).

To sum up, we observe that in the locally Lipschitz non-smooth case it appears that we need different tools (namely, the directional derivative and the Clarke derivatives) to solve problems similar to the ones stated in Section 2 for a smooth function. Thus there is no flexible tool such as the gradient in the Nonsmooth case.

It is worth noting that all the above results are applicable only if one is able to compute numerically the mentioned function values and gradients. In the sequel we shall discuss the possibility of constructing a tool which answers the same questions as the Clarke subdifferential does and which is in some sense simpler.

5. The Michel–Penot Subdifferential

Michel and Penot proposed the following generalized derivative (see [12])

$$f_{\text{mp}}^{\uparrow}(x,g) = \sup_{q\in\mathbb{R}^n}\left\{\limsup_{\alpha\downarrow 0}\frac{1}{\alpha}[f(x+\alpha(g+q)) - f(x+\alpha q)]\right\}. \tag{22}$$

We shall call it the Michel–Penot upper derivative of f at x in the direction g. The quantity

$$f_{\text{mp}}^{\downarrow}(x,g) = \inf_{q\in\mathbb{R}^n}\left\{\liminf_{\alpha\downarrow 0}\frac{1}{\alpha}[f(x+\alpha(g+q) - f(x+\alpha q)]\right\} \tag{23}$$

will be referred to as the Michel–Penot lower derivative of f at x in the direction g. If f is locally Lipschits then there exists a convex compact set $\partial_{\text{mp}}f(x)$ such that

$$f_{\text{mp}}^{\uparrow}(x,g) = \max_{v\in\partial_{\text{mp}}f(x)}(v,g), \tag{24}$$

$$f_{\text{mp}}^{\downarrow}(x,g) = \min_{w\in\partial_{\text{mp}}f(x)}(w,g). \tag{25}$$

Recall that if f is also d.d. then $\partial_{\text{mp}}f(x)$ is the Clarke subdifferential of the function $h(g) = f'(x,g)$ at $g = 0_n$. The set $\partial_{\text{mp}}f(x)$ is often referred to as the small subdifferential. It is known in general that

$$\partial_{\text{mp}}f(x) \subset \partial_{cl}f(x) \tag{26}$$

and in some cases

$$\partial_{\text{mp}}f(x) \neq \partial_{cl}f(x). \tag{27}$$

At the same time, it preserves some of the properties of $\partial_{cl}f(x)$. Namely, the condition

$$O_n \in \partial_{\text{mp}}f(x) \tag{28}$$

is a necessary condition for an extremum (a minimum and a maximum alike). The condition (28) is stronger than the condition that $0 \in \partial_{cl}f(x)$ if (28) holds.

If f is d.d. and (28) is not yet satisfied then there exist directions g's such that

$$f_{\text{mp}}^{\uparrow}(x,g) < 0, \tag{29}$$

and for such a direction one has

$$f'(x,g) \leq f_{\text{mp}}^{\uparrow}(x,g) < 0, \tag{30}$$

while

$$f'(x,-g) \geq f_{\text{mp}}^{\downarrow}(x,-g) = -f_{\text{mp}}^{\downarrow}(x,g) > 0, \tag{31}$$

i.e. all the directions satisfying (29) are descent directions, while the opposite directions are ascent ones. Unfortunately, the robustness of a direction satisfying (29) and (30) is not guaranteed. This is because the mapping $\partial_{\text{mp}} f$ (unlike the mapping $\partial_{cl} f$) is not upper semicontinuous in X.

Hence, the small subdifferential of Michel and Penot has properties similar to te ones of the Clarke subdifferential. Since $\partial_{\text{mp}} f(x)$ is "smaller" than $\partial_{cl} f(x)$, it has some preferences: (i) the necessary condition (28) is stronger; (ii) the family of directions satisfying simultaneously (30) and (31) is in general greater than the similar family obtained by means of the Clarke subdifferential.

Another advantage of the small subdifferential is the fact that if f is quasidifferentiable, one can construct Calculus for computing "small" subdifferentials.

The notion of small subdifferential has been used by many authors to get new results (see [2, 17, 9]). Most generalizations of the concept of gradient are also employed to get new "mean-value theorems" (see [11, 18, 1]).

6. Quasidifferentiable Functions: Calculus for the Michel–Penot Subdifferentials

Let the function f be Lipschitz on an open set $\Omega \subset \mathbb{R}^n$ and quasidifferentiable at $x \in \Omega$, i.e. it is directionally differentiable at x and there exists a pair of convex compact sets $\underline{\partial} f(x), \overline{\partial} f(x) \subset \mathbb{R}^n$ such that, for each $g \in \mathbb{R}^n$

$$f'(x,g) = \max_{v \in \underline{\partial} f(x)} (v,g) + \min_{w \in \overline{\partial} f(x)} (w,g). \tag{32}$$

The pair of sets $\mathcal{D} f(x) = [\underline{\partial} f(x), \overline{\partial} f(x)]$ is called a quasidifferential of f at x. There exists a well-developed Calculus for computing quasidifferentials (see [6,7]) which is a generalization of the classical "smooth" Differential Calculus.

It is convenient to describe necessary optimality conditions in terms of quasidifferentials. For example, a necessary condition for a minimum takes the form

$$-\overline{\partial} f(x^*) \subset \underline{\partial} f(x^*) \tag{32'}$$

and a necessary condition for a maximum is

$$-\underline{\partial} f(x^{**}) \subset \overline{\partial} f(x^{**}). \tag{32"}$$

Let $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^n$ be convex compact sets and let $\rho_A(g)$ and $\rho_B(g)$ be their support functions

$$\rho_A(g) = \max_{v \in A} (v,g), \ \rho_B(g) = \max_{v \in B} (v,g).$$

These functions are almost everywhere differentiable. By definition

$$A \dot{-} B = \text{cl co} \{\nabla \rho_A(g) - \nabla \rho_B(g) | g \in T\} \tag{33}$$

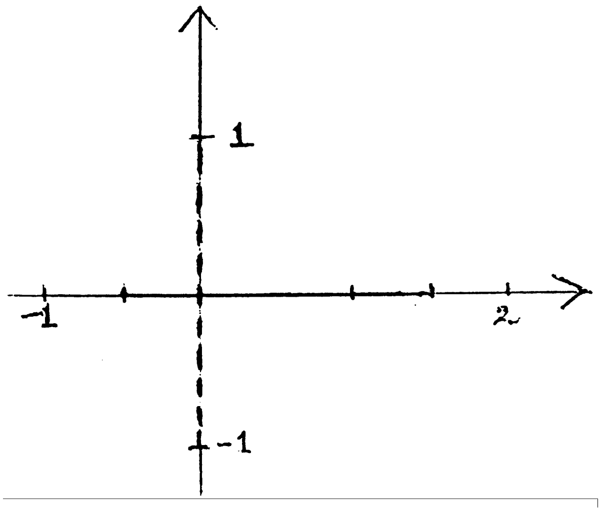

Figure 1.

where T is any set of full measure of differentiability points of ρ_A and ρ_B. The difference of sets just defined is consistent: If

$$A = B + C \quad \text{then} \quad A \dot{-} B = C.$$

As a corollary, we get $A \dot{-} A = O_n$.

The difference $\dot{-}$ was introduced in [5] and studied by Rubinov and Akhundov [17]. It was shown in [17, 7] that

$$\partial_{cl} h(0_n) = \underline{\partial} f(x) \dot{-} (-\overline{\partial} f(x)) \tag{34}$$

where $h(g) = f'(x, g)$.

It was demonstrated in [11,12] that if f is locally Lipschitz and directionally differentiable then

$$\partial_{\text{mp}} f(x) = \partial_{cl} h(0_n). \tag{35}$$

Combining (34) and (35) we arrive at the following result.

THEOREM 6.1. *If f is locally Lipschitz and quasidifferentiable at a point x, then*

$$\partial_{\text{mp}} f(x) = \underline{\partial} f(x) \dot{-} (-\overline{\partial} f(x)). \tag{36}$$

REMARK. Since there exists a Calculus for quasidifferentiable functions, there exists a Calculus for the Michel–Penot subdifferentials: first find a quasidifferential of f at x and then apply Theorem 6.1. (of course, f is Locally Lipschitz and quasidifferentiable at x).

The connection between the Clarke subdifferential and the quasidifferential is discussed in [7, sect.4, ch.3].

7. Convexificators of a Positively Homogeneous Function

Let $h : \mathbb{R}^n \to \mathbb{R}$ be a positively homogeneous (p.h.) (of degree 1) function. We say that a convex compact set $C \subset \mathbb{R}^n$ is a convexificator (CF) of h if

$$\min_{w\in C}(w,g) \leq h(g) \leq \max_{v\in C}(v,g), \quad \forall\, g \in \mathbb{R}^n. \tag{37}$$

Note that max and min in (37) are taken over the same set C!

If h is bounded on $S_1 = \{g \in \mathbb{R}^n \,|\, \|g\| = 1\}$ then there is no problem of existence of a convexificator: any ball centered at the origin with the radius sufficiently large can be taken as a convexificator. The notion of convexificator was introduced and studied in [8] (see also [7]). Note that in many applications the function h is often taken as some kind of a generalized directional derivative.

THEOREM 7.1. *The following properties hold:*

1. *If*

$$h(g) \geq 0 \quad \forall\, g \in \mathbb{R}^n \tag{38}$$

then $0_n \in C$.

2. *If*

$$h(g) > 0 \quad \forall\, g \neq 0_n \tag{39}$$

then

$$0_n \in \operatorname{int} C \tag{40}$$

3. *If*

$$h(g) \leq 0 \quad g \in \mathbb{R}^n \tag{41}$$

then

$$0_n \in C \tag{42}$$

4. *If*

$$h(g) < 0 \quad \forall\, g \neq 0_n \tag{43}$$

then

$$0_n \in \operatorname{int} C \tag{44}$$

5. *If*

$$0_n \notin C \tag{45}$$

then the sets

$$Q^- = \left\{ g \in S_1 \,\middle|\, \max_{v\in C}(v,g) < 0 \right\} \tag{46}$$

$$Q^+ = \left\{ g \in S_1 \middle| \min_{v \in C} (v, g) > 0 \right\} \tag{47}$$

are nonempty and

$$Q^- = -Q^+. \tag{48}$$

Proof.

1. If (38) holds then (37) implies

$$\max_{v \in C} (v, y) \geq 0 \quad \forall\, g \in \mathbb{R}^n. \tag{49}$$

The inequality (49) is equivalent (see [4,6]) to $0_n \in C$.

2. The inequalities (39) and (37) imply

$$\max_{v \in C} (v, g) > 0 \quad \forall\, g \neq 0_n. \tag{50}$$

The inequality (50) is equivalent (see [4,6]) to the inclusion

$$0_n \in \text{ int } C. \tag{51}$$

3. Analogeously, the inequalities (41) and (37) imply

$$\min_{w \in C} (w, g) \leq 0, \quad \forall\, g \in \mathbb{R}^n \tag{52}$$

which is equivalent to

$$\max_{v \in C} (v, g) \geq 0, \quad \forall\, g \in \mathbb{R}^n, \tag{53}$$

and, hence, $0_n \in C$.

4. The inequalities (43) and (37) imply the inequality $\min_{w \in C} (w, g) < 0, \quad \forall\, g \neq 0_n$ which is the same as $\max_{v \in C} (v, g) > 0 \quad \forall\, g \neq 0_n$, and, therefore, $0_n \in \text{ int } C$.

5. If (45) holds then

$$\|v^*\| = \min_{v \in C} \|v\| > 0 \tag{54}$$

It then follows from necessary conditions for a minimum that

$$\max_{v \in C} (v, g_0) \leq -\|v^*\| < 0, \tag{55}$$

where

$$g_0 = -\frac{v^*}{\|v *\|}. \tag{56}$$

Thus, the set Q^- is not empty. The inequality (55) also yields

$$\min_{w \in C} (w, g_1) \geq \|v^*\| > 0 \tag{57}$$

with $g_1 = -g_0$. Hence, the set Q^- is also not empty. It is clear from the definitions (49) and (50) that $Q^+ = -Q^-$.

COROLLARY 7.1. *If (45) holds then (37) implies*

$$h(g) < 0, \quad \forall\, g \in Q^-, \tag{58}$$

$$h(g) > 0, \quad \forall\, g \in Q^+, \tag{59}$$

and (59) can be rewritten in the form

$$h(-g) > 0 \quad \forall\, g \in Q^-. \tag{60}$$

REMARK. The inclusion $0_n \in C$ doesn't necessarily imply that one of the inequalities (38) or (41) holds, as well as the inclusion $0_n \in \text{ int } C$ doesn't necessarily mean that one of the inequalities (39) or (43) holds. However, if h is in addition, convex then the condition $0 \in C$ becomes a necessary and sufficient condition for the inequality

$$h(g) \geq 0 \quad \forall\, g \in \mathbb{R}^n.$$

8. Minimal Convexificators

It follows from definition (37) that a convexificator is not uniquely defined: If C is (CF) of h then the set $C_1 = A + C$ (with A an arbitrary convex compact set) is also a (CF) of h.

We say that C^* is a *minimal convexificator* of h if there exists no other convex compact set C such that $C \subset C^*$, $C \neq C^*$ and C is a (CF) of h.

EXAMPLE 8.1. Let $h : \mathbb{R}^n \to \mathbb{R}$ be a convex positively homogeneous function, then there exists a (unique) convex compact set $A \subset \mathbb{R}^n$ such that

$$h(g) = \max_{v \in A}(v, g). \tag{61}$$

Clearly A is a CF of h. This CF is a minimal one (and unique!).

EXAMPLE 8.2. Let $h : \mathbb{R}^n \to \mathbb{R}$ be a concave positively homogeneous function, then there exists a (unique) convex compact set $B \subset \mathbb{R}^n$ such that

$$h(g) = \min_{w \in B}(w, g). \tag{62}$$

Clearly, B is a (CF). It is a minimal one (and unique!).

EXAMPLE 8.3. Let $h : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$h(g) = h(g_1, g_2) = |g_1| - |g_2|.$$

It is not difficult to check that the set $C = \text{co}\,\{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$ is a (CF):

$$\min_{w \in C}(w, g) \leq h(g) \leq \max_{v \in C}(v, g).$$

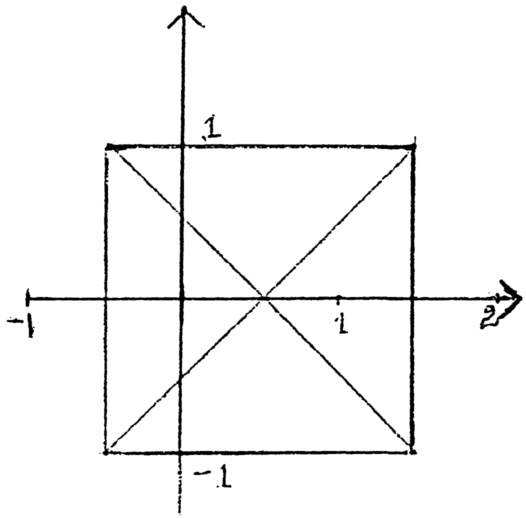

Figure 2.

Take now the sets

$$C_1 = \text{co}\{(1,1),(-1,-1)\}, \quad C_2 = \text{co}\{(-1,1),(1,-1)\}.$$

Direct calculations show that

$$\min_{w \in C_1}(w,g) \leq h(g) \leq \max_{v \in C_1}(v,g)$$

$$\min_{w \in C_2}(w,g) \leq h(g) \leq \max_{v \in C_2}(v,g),$$

i.e. the sets C_1 and C_2 are convexificators as well, and since $C_1 \subset C, C_2 \subset C$ we conclude that C is not a minimal CF, however, the sets C_1 and C_2 are both minimal.

Now one can develop a Calculus for computing convexificators. For example, the following property holds.

THEOREM 8.1. *If functions h_1, h_2 are p.h., $\lambda_1, \lambda_2 \in \mathbb{R}$, C_1, C_2 are convexificators of h_1, h_2, respectively, then the function $h = \lambda_1 h_1 + \lambda_2 h_2$ is also p.h. and the set $C = \lambda_1 C_1 + \lambda_2 C_2$ is a convexificator of h.*

Proof. It suffices to prove the required property for functions $h = h_1 + h_2$ and $h = \lambda h_1$. If h_1, h_2 are positively homogeneous functions then the function $h = h_1 + h_2$ is also p.h. and if C_1, C_2 are convexificators of h_1 and h_2 respectively, then the set $C = C_1 + C_2 = \{v = v_1 + v_2 | v_1 \in C_1, v_2 \in C_2\}$ (the Minkowski sum) is a (CF) of h. Indeed, we have

$$\min_{w \in C_1}(w,g) \leq h_1(g) \leq \max_{v \in C_1}(v,g), \ \forall g \in \mathbb{R}^n \tag{63}$$

$$\min_{w \in C_2} (w, g) \leq h_2(g) \leq \max_{v \in C_2} (v, g), \ \forall g \in \mathbb{R}^n. \tag{64}$$

It follows from (63) and (64) that

$$\min_{w \in C_1} (w, g) + \min_{w \in C_2} (w, g) \leq h_1(g) + h_2(g) \leq \max_{v \in C_1} (v, g) + \max_{v \in C_2} (v, g). \tag{65}$$

Since

$$\min_{w \in C_1} (w, g) + \min_{w \in C_2} (w, g) = \min_{w \in C_1 + C_2} (w, g) = \min_{w \in C} (w, g),$$

$$\max_{v \in C_1} (v, g) + \max_{v \in C_2} (v, g) = \max_{v \in C_1 + C_2} (v, g) = \max_{v \in C} (v, g),$$

the inequalities (65) yield the required property.

Now let h_1 be p.h, $\lambda \in \mathbb{R}, C_1$ be a (CF) of h_1. The function $h = \lambda h_1$ is p.h. for any $\lambda \in \mathbb{R}$. We have

$$\min_{w \in C_1} (w, g) \leq h_1(g) \leq \max_{v \in C_1} (v, g). \tag{66}$$

If $\lambda \geq 0$ then (66) implies

$$\lambda \min_{w \in C_1} (w, g) \leq \lambda h_1(g) \leq \lambda \max_{v \in C_1} (v, g). \tag{67}$$

Since

$$\lambda \min_{w \in C_1} (w, g) = \min_{w \in (\lambda C_1)} (w, g),$$

$$\lambda \max_{v \in C_1} (v, g) = \max_{v \in (\lambda C_1)} (v, g),$$

(67) yields

$$\min_{w \in C} (w, g) \leq h(g) \leq \max_{v \in C} (v, g), \tag{68}$$

where

$$C = \lambda C_1 = \{v = \lambda v_1 | v_1 \in C\}.$$

If $\lambda < 0$ then we get from (66)

$$\lambda \min_{w \in C_1} (w, g) \geq \lambda h_1(g) \geq \lambda \max_{v \in C_1} (v, g). \tag{69}$$

Since $\lambda < 0$

$$\lambda \min_{w \in C_1} (w, g) = \max_{v \in (\lambda C_1)} (v, g), \qquad \lambda \max_{v \in C_1} (v, g) = \min_{w \in (\lambda C_1)} (w, g),$$

therefore (69) again yields

$$\min_{w\in C}(w,g) \le h(g) \le \max_{v\in C}(v,g)$$

with $C = \lambda C_1$.

REMARK. If C_1, C_2 are minimal (CF) of the functions h_1, h_2, respectively, the set $C = C_1 + C_2$ is not necessarily a minimal (CF) of the function $h = h_1 + h_2$.

EXAMPLE 8.4. Let $h_1, h_2 : \mathbb{R}^2 \to \mathbb{R}$ be defined as follows

$$h_1(g) = |g_1|, h_2(g) = -|g_2|.$$

The function h_1 is convex, and the set $C_1 = \text{co}\,\{(1,0),(-1,0)\}$ is a minimal (and unique!) (CF) of h_1 The function h_2 is concave and the set $C_2 = \text{co}\,\{(0,1),(0,-1)\}$ is its minimal (and unique!) (CF). The set

$$C = C_1 + C_2 = \text{co}\,\{(1,1),(1,-1),(-1,1),(-1,1)\}$$

is a (CF) of the function $h = h_1 + h_2$ but not a minimal one (see Example 8.3).

9. Convexificators for Locally Lipschitz Functions

Let the function $f : \Omega \to \mathbb{R}$ be locally Lipschitz on an open set $\Omega \subset \mathbb{R}^n$. Then its upper and lower Dini derivatives

$$f_{\mathcal{D}}^{\uparrow}(x,g) = \limsup_{\alpha\downarrow 0} \frac{f(x+\alpha g) - f(x)}{\alpha}$$

and

$$f_{\mathcal{D}}^{\downarrow}(x,g) = \liminf_{\alpha\downarrow 0} \frac{f(x+\alpha g) - f(x)}{\alpha}$$

are bounded and continuous in g. By the definition (see (16), (17)) we get

$$f_{cl}^{\downarrow}(x,g) \le f_{\mathcal{D}}^{\downarrow}(x,g) \le f_{\mathcal{D}}^{\uparrow}(x,g) \le f_{cl}^{\uparrow}(x,g). \tag{70}$$

Using (18) and (19) we get

$$\min_{w\in\partial_{cl} f(x)}(w,g) \le f_{\mathcal{D}}^{\downarrow}(x,g) \le f_{\mathcal{D}}^{\uparrow}(x,g) \le \max_{v\in\partial_{cl} f(x)}(v,g). \tag{71}$$

Hence, the Clarke subdifferential of f at x is a convexificator of both functions

$$h(g) = f_{\mathcal{D}}^{\uparrow}(x,g) \text{ and } h(g) = f_{\mathcal{D}}^{\downarrow}(x,g).$$

From the definitions (23) and (24) and the formulas (25) and (26) we conclude that

$$\min_{w\in\partial_{\text{mp}}f(x)}(w,g) \le f_{\mathcal{D}}^{\downarrow}(x,g) \le f_{\mathcal{D}}^{\uparrow}(x,g) \le \max_{v\in\partial_{\text{mp}}f(x)}(v,g). \tag{72}$$

Thus, the Michel–Penot subdifferential of f at x is also a convexificator of both functions

$$h(g) = f_{\mathcal{D}}^{\uparrow}(x,g) \text{ and } h(g) = f_{\mathcal{D}}^{\downarrow}(x,g).$$

Since (see (27)) $\partial_{\text{mp}}f(x) \subset \partial_{cl}f(x)$, the Michel–Penot subdifferential of f at x is a "smaller" convexificator of the functions $f_{\mathcal{D}}^{\uparrow}(x,g)$ and $f_{\mathcal{D}}^{\downarrow}(x,g)$ than the Clarke subdifferential.

We shall call a convexificator $C^+(x)$ $(C^-(x))$ of the function $h(g) = f_{\mathcal{D}}^{\uparrow}(x,g)$ $(h(g) = f_{\mathcal{D}}^{\downarrow}(x,g))$ an upper (a lower) convexificator of f at x. If $C(x)$ is a convexificator of both functions $f_{\mathcal{D}}^{\uparrow}(x,g)$ and $f_{\mathcal{D}}^{\downarrow}(x,g)$ we say that $C(x)$ is a convexificator of f at x. Hence, the Clarke and Michel–Penot subdifferentials are both convexificaters of f at x.

Now we are in a position to define a (CF)-mapping for a locally Lipschitz function. We say that a mapping $C^+(C^-) : \Omega \to 2^{\mathbb{R}^n}$ is an upper (a lower) (CF)-mapping of f on Ω if for every $x \in \Omega$ the set $C^+(x)(C^-(x))$ satisfies the following inequalities

$$\min_{w\in C^+(x)}(w,g) \le f_{\mathcal{D}}^{\uparrow}(x,g) \le \max_{v\in C^+(x)}(v,g) \quad \forall\, g \in \mathbb{R}^n \tag{73}$$

$$\left(\min_{w\in C^-(x)}(w,g) \le f_{\mathcal{D}}^{\downarrow}(x,g) \le \max_{v\in C^-(x)}(v,g) \quad \forall g \in \mathbb{R}^n\right).$$

A mapping $C : \Omega \to \mathbb{R}^n$ is called a (CF)-mapping of f is for every $x \in \Omega$ the set $C(x)$ satisfies the inequalities

$$\min_{w\in C(x)}(w,g) \le f_{\mathcal{D}}^{\downarrow}(x,g) \le f_{\mathcal{D}}^{\uparrow}(x,g) \le \max_{v\in C(x)}(v,g) \quad \forall g \in \mathbb{R}^n.$$

If a function f is quasidifferentiable (but not necessarily Lipschitz), at a point x one can construct a convexificator (as $C = \underline{\partial}f(x)\dot{-}(-\overline{\partial}f(x))$) and study the directional derivative by means of this convexificator (or to find a smaller one if possible).

As a corollary of Theorem 7.1 we get the following.

THEOREM 9.1. *Let f be locally Lipschitz on Ω and let C^+ and $C^- : \Omega \to 2^{\mathbb{R}^n}$ be its upper and lower (CF)-mappings. Then the following properties hold:*

1. *For a point x^* to be a minimum point of f it is necessary that*

$$O_n \in C^-(x^*). \tag{74}$$

2. *For a point x^{**} to be a maximum point of f it is necessary that*

$$O_n \in C^+(x^{**}). \tag{75}$$

3. *If $C : \Omega \to 2^{\mathbb{R}^n}$ is a (CF)-mapping of f and if $0_n \notin C(x)$ then the sets*

$$Q^-(x) = \left\{ g \in S_1 | \max_{v \in C(x)} (v, g) < 0 \right\}$$

and

$$Q^+(x) = \left\{ g \in S_1 | \min_{v \in C(x)} (v, g) > 0 \right\}$$

are not empty and $Q^+(x) = -Q^-(x)$. Moreover, for every $g \in Q^-(x)$ the following inequalities hold

$$f_{\mathcal{D}}^{\downarrow}(x, g) \leq \max_{v \in C(x)} (v, g) < 0$$

$$f_{\mathcal{D}}^{\uparrow}(x, -g) \geq \min_{w \in C(x)} (w, g) = - \max_{v \in C(x)} (v, g) > 0.$$

COROLLARY 9.1. *If $C(x)$ is a (CF) of f at x and if x is an extremum of f then*

$$0 \in C(x).$$

REMARK. Now, it is clear why it is of great interest to find a convexificator as "small" as possible: (i) the set of stationary points (the point satisfying (74) and (75) will be smaller (and the amount of points suspicious to be extremal points will be smaller) and (ii) if x is not yet a stationary point we are able to find may be a larger set of directions of descent and ascent which possess the property. If f is also d.d. and its $h'(x, g) < 0$ then $h'(x, -g) > 0$.

EXAMPLE 9.1. Let $x = (x_1, x_2) \in \mathbb{R}^2$; $x_0 = (0, 0)$

$$f(x) = |x_1| - |x_2| + \frac{1}{2}x_1.$$

The function f is directionally differentiable and Lipschitz on $\mathbb{R}^2$. It is easy to check that

$$f'(x_0, g) = |g_1| - |g_2| + \frac{1}{2}g_1, \tag{76}$$

where $g = (g_1, g_2)$. The function f is differentiable everywhere except the points where either $x_1 = 0$ or $x_2 = 0$ or both. Therefore by making use of (19) we can get (see Figure 2)

$$\partial_{cl} f(x_0) = \text{co} \left\{ \left(-\frac{1}{2}, 1 \right), \left(-\frac{1}{2}, -1 \right), \left(\frac{3}{2}, 1 \right), \left(-\frac{3}{2}, -1 \right) \right\}.$$

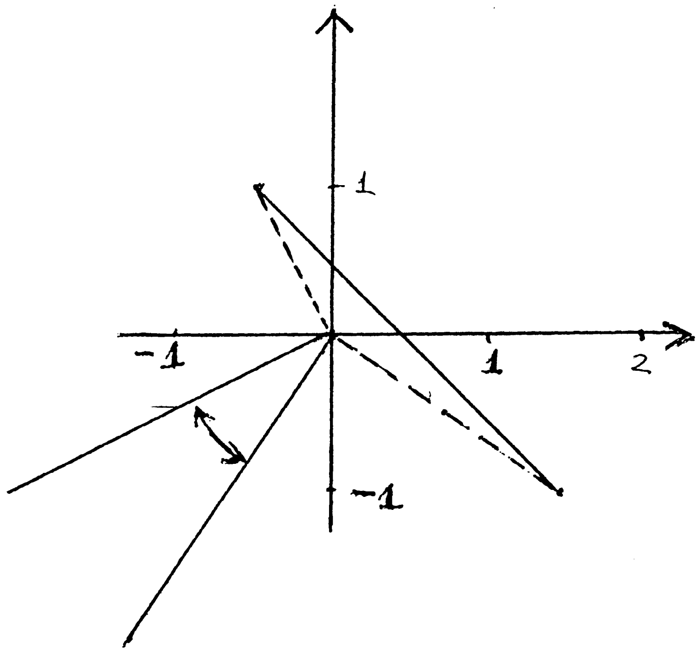

Figure 3.

The function f is also quasidifferentiable on $\mathbb{R}^2$ and one can take the following pair of sets as a quasidifferential of f at x_0:

$$\mathcal{D}f(x_0) = [\underline{\partial} f(x_0), \overline{\partial} f(x_0)],$$

where

$$\underline{\partial} f(x_0) = \text{co}\left\{\left(\frac{3}{2}, 0\right), \left(\frac{-1}{2}, 0\right)\right\},$$
$$\overline{\partial} f(x_0) = \text{co}\left\{(0,1), (0,-1)\right\}.$$

Since (see Figure 1) conditions (32′) and (32") are not satisfied, the point x_0 is neither a minimum point of f, nor its maximum point. At the same time the point x_0 is a Clarke-stationary point since the necessary condition (20) holds at x_0. It follows from (34) and (19) that the function $h(g) = f'(x_0, g)$ has the following Clarke subdifferential at $g_0(0,0)$:

$$\partial_{cl} h(g_0) = \text{co}\left\{\left(-\frac{1}{2}, 1\right), \left(-\frac{1}{2}, -1\right), \left(\frac{3}{2}, 1\right), \left(-\frac{3}{2}, -1\right)\right\}$$

i.e.

$$\partial_{cl} h(g_0) = \partial_{cl} f(x_0).$$

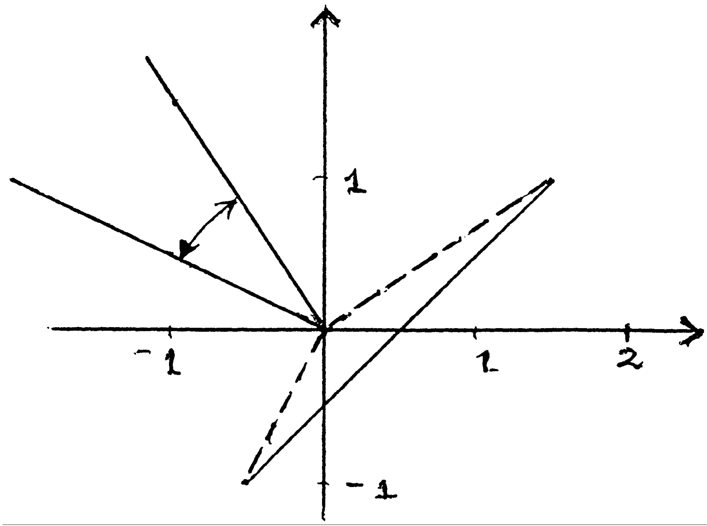

Figure 4.

We can get the same results applying the formula (36):

$$\underline{\partial} f(x_0) \dot{-} (-\overline{\partial} f(x_0) = \text{co}\left\{\left(-\frac{1}{2},1\right),\left(-\frac{1}{2},-1\right),\left(\frac{3}{2},1\right),\left(-\frac{3}{2},-1\right)\right\}.$$

Since $\partial_{mp} f(x_0) = \partial_{cl} h(g_0)$, this implies that

$$\partial_{mp} f(x_0) = \partial_{cl} f(x_0)$$

and the necessary condition (29) holds as well (that is, x_0 is a stationary point in the sense of Michel–Penot).

The set $\partial_{cl} f(x_0) = \partial_{\text{mp}} f(x_0) = \underline{\partial} f(x_0) \dot{-} \overline{\partial} f(x_0)$ is a convexificator of f at x_0. However, it is easy to see that the sets

$$C_1 = \text{co}\left\{\left(-\frac{1}{2},1\right),\left(\frac{3}{2},-1\right)\right\},\ C_2 = \text{co}\left\{\left(-\frac{1}{2},-1\right),\left(\frac{3}{2},1\right)\right\}$$

are also convexificators. Let us check, e.g., that C_1 is a convexificator of f at x_0 :

$$h(g) = f'(x_0, g) = |g_1| - |g_2| + \frac{1}{2} g_1.$$

Note that

$$\min_{v \in C_1} (v, g) \ = \ \min\left\{-\frac{1}{2} g_1 + g_2, \frac{3}{2} g_1 - g_2\right\} \tag{77}$$

$$\max_{v \in C_1} (v, g) \ = \ \max\left\{-\frac{1}{2} g_1 + g_2, \frac{3}{2} g_1 - g_2\right\}. \tag{78}$$

The following four cases are possible

1. $g_1 \geq 0,\ g_2 \geq 0$
2. $g_1 \geq 0,\ g_2 \leq 0$
3. $g_1 \leq 0,\ g_2 \geq 0$
4. $g_1 \leq 0,\ g_2 \leq 0$

In case (1): $h(g) = \frac{3}{2}g_1 - g_2$. Then, clearly

$$\min_{v \in C_1}(v, g) \leq h(g) \leq \max_{v \in C_1}(v, g). \tag{79}$$

In case (2): $h(g) = g_1 + g_2 + \frac{1}{3}g_1 = \frac{3}{2}g_1 + g_2$.

Since $g_1 \geq 0_1, g_2 \leq 0, -\frac{1}{2}g_1 + g_2 \leq 0,\ \frac{3}{2}g_1 - g_2 \geq 0$ and

$$\min_{v \in C_1}(v, g) = -\frac{1}{2}g_1 + g_2 \leq \frac{3}{2}g_1 + g_2 = h(g). \tag{80}$$

Analogously,

$$\max_{v \in C_1}(v, g) = \frac{3}{2}g_1 - g_2 \geq \frac{3}{2}g_1 + g_2 = h(g) \tag{81}$$

(80) and (81) imply again (29).

In case (3): $h(g) = -g_1 - g_2 + \frac{1}{2}g_1 = -\frac{1}{2}g_1 - g_2$.

Since $g_1 \leq 0, g_2 \geq 0, -\frac{1}{2}g_1 + g_2 \geq 0,\ \frac{3}{2}g_1 - g_2 \leq 0$ and so,

$$\min_{v \in C_1}(v, g) = \frac{3}{2}g_1 - g_2 \leq -\frac{1}{2}g_1 - g_2 = h(g) \tag{82}$$

$$\max_{v \in C_1}(v, g) = -\frac{1}{2}g_1 + g_2 \geq -\frac{1}{2}g_1 - g_2 = h(g) \tag{83}$$

(82) and (83) yield (79).

In case (4): $h(g) = -g_1 + g_2 + \frac{1}{2}g_1 = -\frac{1}{2}g_1 + g_2$.

It is clear from (77) and (78) that

$$\min_{v \in C_1}(v, g) \leq -\frac{1}{2}g_1 + g_2 = h(g) \leq \max_{v \in C_1}(v, g). \tag{84}$$

Thus, in all the cases (1)–(4) we get (79), i.e. C_1 is a convexificator.

In a similar way we can prove that C_2 is a convexificator as well. These convexificators are minimal. On the other hand, $C_1 \subset \partial_{mp}f(x_0)$, $C_2 \subset \partial_{mp}f(x_0)$ and $C_1 \neq \partial_{mp}f(x_0)$, $C_2 \neq \partial_{mp}f(x_0)$ that is, these convexificators are "smaller" than $\partial_{mp}f(x_0)$.

Let us take C_1. Since (see Figure 3) $0_2 \notin C_1$, the point x_0 is not a stationary one (and, thus, is "out of the list of suspicious points"). The set Q_1^- (see (49)) consists of unit vectors located between the points $\left(-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right)$ and $\left(-\frac{2}{\sqrt{13}}, -\frac{3}{\sqrt{13}}\right)$ (these two points don't belong to Q_1^-). Using the relation (48), we can get Q_1^+ (see (47)).

Consider now the set C_2. Since (see Figure 4) $0_2 \notin C_2$, the point x_0 is not a stationary. The set Q_2^- for this convexificator is the set of unit vectors located between the points $\left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$ and $\left(-\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}}\right)$.

Thus, using the convexificators C_1 and C_2 we conclude that the set $Q^- = Q_1^- \cup Q_2^-$ has the property:

$$h(g) < 0,\ h(-g) > 0 \quad \forall\, g \in Q^-.$$

THEOREM 9.2. (Mean-Value Theorem). *Let C^+ and $C^- : \Omega \to 2^{\mathbb{R}^n}$ be an upper and a lower (CF)-mappings of f, respectively. If the interval co $\{x_1, x_2\} \subset \Omega$ then there exists an $\gamma \in (0,1)$ such that at least one of the following statements hold*

1. *there exists $v \in C^+(x_1 + \gamma(x_2 - x_1))$ such that*

$$f(x_2) - f(x_1) = (v, x_2 - x_1) \tag{85}$$

2. *there exists $v \in C^-(x_1 + \gamma(x_2 - x_1))$ such that (85) is valid.*

Proof. Let us consider the function

$$h(\alpha) = f(x_1 + \alpha(x_2 - x_1)) - f(x_1) + \alpha[f(x_1) - f(x_2)].$$

Since $h(0) = h(1) = 0$, there exists a $\gamma \in (0,1)$ such that the function h attains its extremal value on the interval [0,1] at γ. Let, for example, γ be a minimum point of h. Then the necessary condition for a minimum

$$h_{\mathcal{D}}^{\downarrow}(\gamma, g) \geq 0 \quad \forall\, g \in \mathbb{R} \tag{86}$$

holds, where

$$\begin{aligned} h_{\mathcal{D}}^{\downarrow}(\gamma, g) &= \liminf_{\beta \downarrow 0} \frac{h(\gamma + \beta g) - h(\gamma)}{\beta} \\ &= \liminf_{\beta \downarrow 0} \frac{1}{\beta}[f(x_1 + (\gamma + \beta g)(x_2 - x_1)) - f(x_1 + \gamma(x_2 - x_1)) \\ &\quad +\beta g(f(x_1) - f(x_2))] = f_{\mathcal{D}}^{\downarrow}(x_1 + \gamma(x_2 - x_1), g(x_2 - x_1)) \\ &\quad +g[f(x_1) - f(x_1)]. \end{aligned} \tag{87}$$

Let $x_2 = x_1 + \gamma(x_2 - x_1), \Delta = x_2 - x_1$.

If $g = +1$ then

$$f^{\downarrow}_{\mathcal{D}}(x_\gamma, \Delta) + f(x_1) - f(x_2) \geq 0. \tag{88}$$

If $g = -1$ then

$$f^{\downarrow}_{\mathcal{D}}(x_\gamma, -\Delta) - [f(x_1) - f(x_2)] \geq 0. \tag{89}$$

The inequalities (88) and (89) yield

$$f^{\downarrow}_{\mathcal{D}}(x_\gamma, -\Delta) \leq f(x_2) - f(x_1) \leq f^{\downarrow}_{\mathcal{D}}(x_\gamma, \Delta). \tag{90}$$

By the definition of a convexificator

$$\min_{w \in C^-(\gamma)} (w, g) \leq f^{\downarrow}_{\mathcal{D}}(x_\gamma, g) \leq \max_{v \in C^-(\gamma)} (v, g), \tag{91}$$

where $C^-(\gamma) = C^-(x_1 + \gamma(x_2 - x_1))$.

It follows from (90) and (91) that

$$\min_{v \in C^-(\gamma)} (v, x_2 - x_1) \leq f(x_2) - f(x_1) \leq \max_{v \in C^-(\gamma)} (v, x_2 - x_1). \tag{92}$$

Since the function $(v, x_2 - x_1)$ is continuous in v, it follows from (92) that for some $v \in C^-(\gamma)$ we get $f(x_2) - f(x_2) = (v, x_2 - x_1)$.

Similarly, if γ is a maximizer of h we employ the necessary condition for a maximum:

$$h^{\uparrow}_{\mathcal{D}}(\gamma, g) \geq 0 \quad \forall\, g \in \mathbb{R}$$

and follow the same line of arguments as above.

COROLLARY 9.2. *If* $C : \Omega \to 2^{\mathbb{R}^n}$ *is a* (CF) *mapping of* f *then there exist a* $\gamma \in (0, 1)$ *and* $v \in C(x_1 + \gamma(x_2 - x_1))$ *such that (85) holds*

This result also follows from Theorem 3.1 of [17] in the locally Lipschitz case.

REMARK. As corollaries we get the Clarke mean-value theorem [1] and the Michel–Penot mean-value theorem [12].

For related mean-value theorems which employ the Clarke and Michel–Penot subdifferentials see [17]. The main-value theorem of [17] was proved using a two-sided upper convex approximation which differs from the upper and lower convexicators, examined here.

References

1. Clarke, F.M. (1983), *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York.
2. Crave, B.D., Ralph, D. and Glover, B.M. (1995), Small convex-valued subdifferentials in mathematical programming, *Optimization*, **33**, pp.1–21.

3. Danskin, J. (1967) *Theory of Max-Min and Its Application to Weapons Allocation Problems*, Springer-Verlag, New York.
4. Demyanov, V.F. and Malozemov, V.N. (1994), *Introduction to Minimax*, Wiley Interscience, New York.
5. Demyanov, V.F. (1980) On relationship between the Clarke subdifferential and the quasidifferential, *Vestnik of Leningrad University* **13**, pp. 18–24.
6. Demyanov, V.F. and Vasiliev, L.V. (1986), *Nondifferentiable Optimization*, Springer-Optimization Software, New York.
7. Demyanov, V.F. and Rubinov, A.M. (1995), *Constructive Nonsmooth Analysis*, Verlag Peter Lang, Frankfurt a/M
8. Demyanov, V.F. (1994), Convexification and concavification of a positively homogeneous function by the same family of linear functions, Università di Pisa, Report 3,208,802.
9. Ioffe, A.D. (1993), A Lagrange multiplier rule with small convex-valued subdifferentials for nonsmooth problems of mathematical programming involving equality and non-functional constraints, *Math. Progr.*, **58**, pp. 137–145.
10. Jeyakumar, V. (1987), On optimality conditions in nonsmooth inequality constrained minimization, *Numer. Funct. Anal. and Optim.*, **9** (5 & 6), pp. 536–546.
11. Kuntz, L. and Scholtes, S. (1993), Constraint qualifications in quasidifferentiable optimization, *Mathematical Programming*, **60**, pp. 339–347.
12. Michel. P. and Penot, J.-P. (1984), Calcus sous-differential pour les fonctions lipschitziennes et non-lipschitziennes, *C.R. Acad. Sc. Paris, Ser.I* **298**, pp. 269–272.
13. Pschenichniy, B.N. (1980), *Convex Analysis and Extremal Problems*, Nauka, Moscow.
14. Rubinov, A.M. and Akhundov, I.S. (1992), Difference of compact sets in the sense of Demyanov and its application to nonsmooth analysis, *Optimization*, **23**, pp. 179–188.
15. Treiman, J.S. (1988), Shrinking generalized gradients, *Nonlinear Analysis TM & A.*, **12**, pp. 1429–1449.
16. Rockafellar, R.T. (1970), *Convex Analysis*, Princeton University Press, Princeton N.J.
17. Studniarski, M. and Jeyakumar, V. (1993), A generalized mean-value theorem and optimality conditions in composite nonsmooth minimization, *Nonlinear Analysis TM & A*, **24**(6), 883–894.
18. Xia, Z.-Q. (1987), On mean-value theorems in quasidifferential Calculus, *J. Math. Res. and Exposition, D.I.T., Dalian (China)* **4,** pp. 681–684.