Journal of Convex Analysis
Volume 4 (1997), No. 2, 221–234

Proto-Derivatives of Partial Subgradient Mappings*

R. A. Poliquin
Dept. of Mathematical Sciences, Univ. of Alberta,
Edmonton, Alberta, T6G 2G1, Canada.
rene.poliquin@ualberta.ca

R. T. Rockafellar
Dept. of Mathematics, Univ. of Washington,
Seattle, WA 98195, USA.
rtr@math.washington.edu

Received September 24, 1996
Revised manuscript received May 28, 1997

Partial subgradient mappings have a key role in the sensitivity analysis of first-order conditions for optimality, and their generalized derivatives are especially important in that respect. It is known that such a mapping is proto-differentiable when it comes from a fully amenable function with compatible parameterization, which is a common case in applications; the proto-derivatives can be evaluated then through projections. Here this result is extended to a still broader class of functions than fully amenable, namely, ones obtained by composing a C^2 mapping with a kind of piecewise-C^2 convex function under a constraint qualification.

Keywords: Variational analysis, subgradient mappings, proto-derivatives, second-order epi-derivatives, amenable functions, piecewise-C^2 functions, nonsmooth analysis

1991 Mathematics Subject Classification: Primary 49A52, 58C06, 58C20; Secondary 90C30

1. Introduction

Under natural assumptions, many common set-valued mappings in optimization and variational analysis are *proto-differentiable*. For example, the mapping that gives the feasible set in a parametrized optimization problem is typically proto-differentiable, and the same for the mapping that gives the solution set to a parameterized variational inequality, as shown in [1]. Proto-derivatives were introduced in that paper and further studied in [2]–[16]. They arise from graphical geometry in the following manner.

* This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under grant OGP41983 for the first author and by the National Science Foundation under grant DMS–9500957 for the second author. This paper was written in part while the first author was visiting the Université de Perpignan; thanks to D. Azé, J.-N. Corvellec, C.D. Horvath and the department of mathematics.

ISSN 0944-6532 / \$ 2.50

Let $\Gamma : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ be a set-valued mapping, and let $\bar{z} \in \Gamma(\bar{w})$. For $t > 0$, the difference quotient mapping at $\bar{w}$ for $\bar{z}$ is

$$(\Delta_t \Gamma)_{\bar{w},\bar{z}} : \omega \mapsto [\Gamma(\bar{w} + t\omega) - \bar{z}]/t.$$

If, as $t \searrow 0$, the graph of $(\Delta_t \Gamma)_{\bar{w},\bar{z}}$ in $\mathbb{R}^d \times \mathbb{R}^n$ converges (in the Painlevé-Kuratowski sense) to some set G, one says that Γ is *proto-differentiable* at $\bar{w}$ for $\bar{z}$. The limit mapping having G as its graph is then the *proto-derivative* of Γ at $\bar{w}$ for $\bar{z}$ and is denoted by $\Gamma'_{\bar{w},\bar{z}}$. It associates with each $\omega \in \mathbb{R}^d$ a (sometimes possibly empty) set $\Gamma'_{\bar{w},\bar{z}}(\omega) \subset \mathbb{R}^n$.

This paper centers on the proto-differentiability of an important kind of set-valued mapping in variational analysis: the *partial subgradient* mapping associated with a bivariate function. For a lower semicontinuous, proper function $f : \mathbb{R}^n \to \overline{\mathbb{R}} := [-\infty, \infty]$, the subgradient mapping $\partial f : x \in \mathbb{R}^n \rightrightarrows v \in \mathbb{R}^n$ gives the set of (limiting proximal) subgradients v to f at x, i.e., $v \in \partial f(x)$ provided there exists $x_k \to x$ and $v_k \in \partial_p f(x_k)$ with $f(x_k) \to f(x)$ and $v_k \to v$ (here $\partial_p f(x_k)$ denotes the set of proximal subgradients to f at x_k). When x is written as (x_1, x_2) with $x_1 \in \mathbb{R}^{n_1}$, $x_2 \in \mathbb{R}^{n_2}$ and $n_1 + n_2 = n$, subgradients can be taken in the first argument alone to get the partial subgradient mapping $\partial_1 f : \mathbb{R}^n \rightrightarrows \mathbb{R}^{n_1}$, which assigns to each $x = (x_1, x_2)$ the set $\partial_1 f(x)$ of subgradients of $f(\cdot, x_2)$ at x_1.

Much effort has been devoted to identifying functions that have a proto-differentiable subgradient mapping or partial subgradient mapping. The motivation for such an endeavor comes mainly from sensitivity analysis. The proto-derivative of the inverse subgradient mapping (which exists if and only if the proto-derivative of the subgradient mapping exists), or of the inverse partial subgradient mapping, gives information on the rate of change of "quasi-solutions." These motivations are explained in detail by Levy and Rockafellar [17]. The purpose of this paper is to provide new examples of functions with proto-differentiable partial subgradient mapping.

The concept of a function f being *amenable* at a point $\bar{x}$ has been central to examples of subgradient proto-differentiability. This refers to the existence on some neighborhood of $\bar{x}$ of a composite representation $f(x) = g(F(x))$ in which $F : \mathbb{R}^n \to \mathbb{R}^m$ is a C^1 mapping, $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ is a lower semicontinuous, proper, convex function, and the *basic constraint qualification* holds:

$$\left.\begin{aligned} y \in N_{\operatorname{dom} g}(F(\bar{x})) \\ \nabla F(\bar{x})^* y = 0 \end{aligned}\right\} \quad \Longrightarrow \quad y = 0. \tag{1.1}$$

(Here $\operatorname{dom} g = \{y | g(y) < +\infty\}$, while $\nabla F(\bar{x})$ is the $m \times n$ Jacobian matrix for F at $\bar{x}$, and $*$ denotes transpose.) One says that f is *strongly* amenable when the representation can be chosen with F of class C^2, and that f is *fully* amenable if, in addition, g can be taken to be *piecewise linear-quadratic* (p.l.q.), meaning that $\operatorname{dom} g$ is the nonempty union of finitely many polyhedral (convex) sets, relative to each of which g is quadratic (with affine as a special case).

The class of fully amenable functions is more widespread in variational analysis and optimization than its definition might at first seem to suggest, cf. [7], [8], [15], [16]; its unusually favorable properties have therefore attracted interest. The proto-differentiability of the subgradient mappings ∂f associated with fully amenable functions f was demonstrated by Poliquin [4]. This result has been extended to partial subgradient mappings $\partial_1 f$ in the following way by Levy and Rockafellar [17].

Theorem 1.1. [17] *Suppose for a function $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$ and a point $\bar{x} = (\bar{x}_1, \bar{x}_2)$ that $f(x_1, \bar{x}_2)$ is fully amenable in x_1 at $\bar{x}_1$ with compatible parameterization in x_2 at $\bar{x}_2$. Then for all x sufficiently near to $\bar{x}$ and for all $v_1 \in \partial_1 f(x)$, the partial subgradient mapping $\partial_1 f$ is proto-differentiable at x for v_1, moreover with*

$$\begin{aligned} \partial_1 f(x) &= \operatorname{proj}_1 \partial f(x), \\ (\partial_1 f)'_{x,v_1}(\xi) &= \bigcup\nolimits_v \Big\{ \operatorname{proj}_1 (\partial f)'_{x,v}(\xi) \,\Big|\, \operatorname{proj}_1 v = v_1 \Big\}. \end{aligned}$$

Here proj_1 denotes the projection mapping from $\mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ onto $\mathbb{R}^{n_1}$. The notion of *amenability in x_1 at $\bar{x}_1$ with compatible parameterization in x_2 at $\bar{x}_2$* adapts the constraint qualification in the definition of amenability to the setting of $f(x_1, x_2) = g(F(x_1, x_2))$ by modifying (1.1) to

$$\left.\begin{aligned} y \in N_{\operatorname{dom} g}(F(\bar{x})) \\ \nabla_1 F(\bar{x})^* y = 0 \end{aligned}\right\} \quad \Longrightarrow \quad y = 0, \tag{1.2}$$

where $\nabla_1 F(\bar{x}) = \nabla_1 F(\bar{x}_1, \bar{x}_2)$ is the $m \times n_1$ Jacobian matrix for $F(x_1, \bar{x}_2)$ with respect to x_1 at $\bar{x}_1$. Because of the convexity of $\operatorname{dom} g$, the constraint qualifications (1.1) and (1.2) are equivalent respectively to

$$0 \in \operatorname{int}(\operatorname{dom} g - M) \quad \text{for} \quad M = \{F(\bar{x}) + \nabla F(\bar{x})\xi \mid \xi \in \mathbb{R}^n\}, \tag{1.1$'$}$$

$$0 \in \operatorname{int}(\operatorname{dom} g - M_1) \quad \text{for} \quad M_1 = \{F(\bar{x}) + \nabla_1 F(\bar{x})\xi_1 \mid \xi_1 \in \mathbb{R}^{n_1}\}. \tag{1.2$'$}$$

The proof of Theorem 1.1 in [17] relied heavily on *full* amenability of f and especially on the fact that the subgradient mapping ∂g of the p.l.q. convex function g involved in representing such a function f is piecewise polyhedral: its graph is the union of finitely many polyhedral sets. (This was proved by Sun [18]; see also Rockafellar and Wets [16].) For that reason, the prospects of extending Theorem 1.1 beyond the class of fully amenable functions have seemed slim. But hope has emerged from progress on another front, that of extending the theory of first- and second-order epi-derivatives of a fully amenable functions to some kinds of functions that are only strongly amenable. Although we will not go into that theory here, such epi-derivatives of f were instrumental originally in [4] in getting the proto-differentiability of ∂f. The proof of their existence in [15] hinged on the fact that second-order epi-derivatives of a p.l.q. function g can be expressed through pointwise convergence as well as through epi-convergence in their definition. But Azé and Poliquin [19] have revealed that crucial aspects of the second-order epi-derivatives of p.l.q. functions carry over to the broader class of "piecewise-C^2 functions with polyhedral pieces."

Definition 1.2. A proper function $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ is *piecewise-C^k* if $\operatorname{dom} g$ is the nonempty union of a finite collection of closed sets D_i (for $i = 1, \dots, s$), and on each D_i one has $g(x) = g_i(x)$ for some C^k function defined on an open set that includes D_i. Such a function is said to have *polyhedral pieces* if, in addition, the sets D_i are polyhedral.

Every p.l.q. function g is obviously piecewise-C^2 with polyhedral pieces. Azé and Poliquin showed in [19] that the functions f obtainable through the composition of a C^2 mapping F with a convex function g that is piecewise-C^2 with polyhedral pieces, under the basic constraint qualification (1.1), are twice epi-differentiable. The subgradient mapping ∂f for such a strongly amenable function f is proto-differentiable; cf. [19, Corollary 3.12]. This is due to [6, Theorem 2.2] which says that the *primal-lower-nice* function f is twice epi-differentiable at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if and only ∂f is proto-differentiable at $\bar{x}$ for $\bar{v}$; this class of functions includes all strongly amenable functions. (For lsc, proper, convex functions, which are strongly amenable in particular, the equivalence between second-order epi-differentiability of f and proto-differentiability of ∂f was established earlier in [2].) This result was recently extended to the larger class of *prox-regular* functions in Poliquin and Rockafellar [12, Theorem 6.1].

These developments, while encouraging, have not answered the question of what might be said about *partial* subgradient mappings of functions that are strongly amenable but not fully amenable. Our main result lies in this direction. We state it after another definition, which will be needed in the theorem's hypothesis.

Definition 1.3. A set $D \subset \mathbb{R}^m$ will be said to have a *regular C^2 constraint representation* at a point $\bar{u} \in D$ if there is a neighborhood U of $\bar{u}$ along with C^2 functions h_j for $j = 1, \ldots, r$ and an integer $q \in [0, r]$ such that $D \cap U$ is the set of points $u \in U$ such that

$$h_j(u) - h_j(\bar{u}) \begin{cases} \leq 0 & \text{for } j = 1, \ldots, q, \\ = 0 & \text{for } j = q+1, \ldots, r, \end{cases}$$

and this system satisfies the Mangasarian-Fromovitz constraint qualification at $\bar{u}$ i.e., that the only $\mu = (\mu_1, \ldots, \mu_r)$ with $\mu_j \geq 0$ for $j = 1, \ldots, q$ and $\sum_{j=1}^{r} \mu_j \nabla h_j(\bar{u}) = 0$ is $\mu = (0, \ldots, 0)$.

Any polyhedral set has a regular C^2 constraint representation; indeed, the constraint functions can be chosen to be linear.

Theorem 1.4. *Suppose for a function $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \overline{\mathbb{R}}$ and a point $\bar{x} = (\bar{x}_1, \bar{x}_2)$ that $f(x_1, \bar{x}_2)$ is strongly amenable in x_1 at $\bar{x}_1$ with compatible parameterization in x_2 at $\bar{x}_2$, with the convex function g in the associated representation $f = g \circ F$ being piecewise-C^2. Assume further that*

(a) *$\operatorname{dom} g$ has a regular C^2 constraint representation at $F(\bar{x})$;*
(b) *∂f is proto-differentiable at $\bar{x}$ for every $v \in \partial f(\bar{x})$;*

these properties hold in particular when g is piecewise-C^2 with polyhedral pieces.

Then, for all x sufficiently near to $\bar{x}$ and for all $v_1 \in \partial_1 f(x)$, the partial subgradient mapping $\partial_1 f$ is proto-differentiable at x for v_1 with

$$\partial_1 f(x) = \operatorname{proj}_1 \partial f(x) = \nabla_1 F(x)^* \partial g(F(x)), \tag{1.4}$$

$$(\partial_1 f)'_{x,v_1}(\xi) = \bigcup_v \Big\{ \operatorname{proj}_1 (\partial f)'_{x,v}(\xi) \,\Big|\, \operatorname{proj}_1 v = v_1 \Big\}. \tag{1.5}$$

Theorem 1.4 obviously implies Theorem 1.1, since p.l.q. functions are piecewise-C^2 with polyhedral pieces. Theorem 1.4 will be proved in the next section after some preliminaries.

2. Proof Developments

A fact about subgradients of piecewise-C^k convex functions needs to be developed first.

Proposition 2.1. *Let $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ be convex and piecewise-C^1, and let L be the affine hull of the convex set $D = \operatorname{dom} g$. The expression in* Definition 1.2 *can then be chosen so that $D_i = \operatorname{cl}(\operatorname{int}_L D_i)$ for all i, where int_L denotes the interior relative to L. In that case one has at every point $u \in D$ that*

$$\partial g(u) = \operatorname{co}\{\nabla g_i(u) \mid i \in I(u)\} + N_D(u), \quad \textit{where } I(u) = \{i \,|\, u \in D_i\}.$$

Proof. The set D, expressible by Definition 1.2 as the union of finitely many closed sets D_i, is closed. By a translation of g in $\mathbb{R}^m$, we can arrange that $0 \in L$, so that L is a subspace. From convex analysis we know then that $N_D(u) + L^\perp = N_D(u)$ for all $u \in D$. Thus the claimed subgradient formula, along with the proposed improvement in the "pieces" that represent g_i, depends only on the restriction of g to the subspace L. Without loss of generality, therefore, we can reduce to the case where D is m-dimensional, or in other words where $L = \mathbb{R}^m$.
In that setting we invoke the Baire category theorem with respect to sets $B \cap D$ for balls B with $\operatorname{int} D \cap \operatorname{int} B \neq \emptyset$: it is impossible for $B \cap D$ to be expressed as the union of countably many sets that are nowhere dense in $B \cap D$, so at least one of the sets $B \cap D_i$ must meet the interior of $B \cap D$, which by convexity is $\operatorname{int} D \cap \operatorname{int} B$. This tells us that $B \cap D$ must meet $\operatorname{int} D_i$ for at least one i. Let $D_i' = \operatorname{cl}(\operatorname{int} D_i)$ and organize the indices so that the nonempty sets among these are the ones for $i = 1, \ldots, s'$. Then $D = D_1' \cup \cdots \cup D_{s'}'$, with $D_i' \subset D_i$ and $D_i' = \operatorname{cl}(\operatorname{int} D_i')$, and these sets therefore furnish the desired improvement.
Suppose henceforth that the D_i's themselves satisfy $D_i = \operatorname{cl}(\operatorname{int} D_i)$. Every point of D is then the limit of points in the interior of some D_i. Because g is convex, the gradient mapping ∇g is continuous relative to its domain of existence (cf. [20, Thm. 25.5]), which includes the open sets $\operatorname{int} D_i$. In addition we know from the definition of a piecewise-C^1 function, that the gradient mapping for each function g_i is continuous on an open set that includes D_i. Hence, whenever $\nabla g(u)$ exists, we have $\nabla g(u) = \nabla g_i(u)$ for some $i \in I(u)$. The claimed formula is immediate then from the formula for ∂g in terms of ∇g provided in [20, Thm. 25.6]. □

Next we look more closely at proto-differentiation. For a set-valued mapping $\Gamma : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ and elements $\bar{z} \in \Gamma(\bar{w})$, the proto-derivative has been defined by $\operatorname{gph} \Gamma'_{\bar{w},\bar{z}} = \lim_{t \searrow 0} \operatorname{gph} (\Delta_t \Gamma)_{\bar{w},\bar{z}}$, when this limit exists. More generally, we define the mapping $\Gamma'^{+}_{\bar{w},\bar{z}} : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ by $\operatorname{gph} \Gamma'^{+}_{\bar{w},\bar{z}} = \limsup_{t \searrow 0} \operatorname{gph}(\Delta_t \Gamma)_{\bar{w},\bar{z}}$.

Proposition 2.2. *For a multifunction $\Gamma : \mathbb{R}^d \rightrightarrows \mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, let $\operatorname{proj}_1 \Gamma : \mathbb{R}^m \rightrightarrows \mathbb{R}^{n_1}$ assign to each w the image of $\Gamma(w)$ under the projection $\operatorname{proj}_1 : \mathbb{R}^n \to \mathbb{R}^{n_1}$. Let $\bar{z}_1 \in (\operatorname{proj}_1 \Gamma)(\bar{w})$ and assume that whenever $z_1' \in (\operatorname{proj}_1 \Gamma)'^{+}_{\bar{w},\bar{z}_1}(\omega)$ (for some ω) there exist $t_k \searrow 0$, $w_k \to \bar{w}$ with $(w_k - \bar{w})/t_k \to \omega$, $(z_n^1, z_n^2) \in \Gamma(w_k)$, $(\bar{z}_1, \bar{z}_2) \in \Gamma(\bar{w})$ with $(z_k^1 - \bar{z}_1)/t_k$ converging to z_1' and $(z_k^2 - \bar{z}_2)/t_k$ converging to some z_2'. Under these assumptions one has for all $\omega \in \mathbb{R}^d$ that*

$$(\operatorname{proj}_1 \Gamma)'^{+}_{\bar{w},\bar{z}_1}(\omega) \;\subset\; \bigcup\nolimits_{\bar{z}} \Big\{ \operatorname{proj}_1(\Gamma'^{+}_{\bar{w},\bar{z}})(\omega) \;\Big|\; \operatorname{proj}_1 \bar{z} = \bar{z}_1 \Big\}.$$

Proof. Immediate from the definitions. □

Corollary 2.3. *Under the assumptions of* Proposition 2.2, *if* Γ *is proto-differentiable at* $\bar{w}$ *for some* $\bar{z} \in \Gamma(\bar{w})$ *with* $\operatorname{proj}_1 \bar{z} = \bar{z}_1$ *then* $(\operatorname{proj}_1 \Gamma)$ *is proto-differentiable at* $\bar{w}$ *for* $\bar{z}_1$ *and for all* $\omega \in \mathbb{R}^d$ *one has*

$$(\operatorname{proj}_1 \Gamma)'_{\bar{w},\bar{z}_1}(\omega) \;=\; \bigcup\nolimits_{\bar{z}} \Big\{ \operatorname{proj}_1(\Gamma'_{\bar{w},\bar{z}})(\omega) \;\Big|\; \operatorname{proj}_1 \bar{z} = \bar{z}_1 \Big\}.$$

Proof. This follows from Prop. 2.2 and [17, Prop. 2.2]. □

Proof of Theorem 1.4. Because of the constraint qualification built into the amenability assumption, the partial subgradient formula (1.4) holds on the basis of the general calculus in [8]. To obtain the desired proto-differentiability of $\partial_1 f$ and the formula (1.5), we can utilize Corollary 2.3 and concentrate simply on verifying that the assumptions of Proposition 2.2 are fulfilled by $\Gamma = \partial f$. This amounts to producing, for arbitrary $\bar{v}_1 \in \partial_1 f(\bar{x})$, $\xi \in \mathbb{R}^n$, and $v_1' \in (\partial_1 f)'^{+}_{\bar{x},\bar{v}_1}(\xi)$, elements $t_k \searrow 0$, $x_k \to \bar{x}$ and $(v_{1,k}, v_{2,k}) \in \partial f(x_k)$ and $\bar{v}_2$ with $(\bar{v}_1, \bar{v}_2) \in \partial f(\bar{x})$ such that $(x_k - \bar{x})/t_k \to \xi$, $(v_{1,k} - \bar{v}_1)/t_k \to v_1'$ and $(v_{2,k} - \bar{v}_2)/t_k$ converges to some v_2'.
In taking $v_1' \in (\partial_1 f)'^{+}_{\bar{x},\bar{v}_1}(\xi)$, we automatically have by definition the existence of $t_k \searrow 0$, $x_k \to \bar{x}$ and $v_{1,k} \in \partial f(x_k)$ such that $(x_k - \bar{x})/t_k \to \xi$ and $(v_{1,k} - \bar{v}_1)/t_k \to v_1'$. From (1.4), we can find $y_k \in \partial g(F(x_k))$ with $v_{1,k} = \nabla_1 F(x_k)^* y_k$. The set $\operatorname{gph} \partial g$ is closed, and the constraint qualification (1.2) guarantees that the sequence of vectors y_k is bounded, so we may assumewithout loss of generality that $y_k \to \bar{y}$ for a vector $\bar{y} \in \partial g(F(\bar{x}))$. Then $\bar{v}_1 = \nabla_1 F(\bar{x})^* \bar{y}$. Let $\xi_k = (x_k - \bar{x})/t_k$, so that $\xi_k \to \xi$. We have

$$\begin{aligned} \frac{v_{1,k} - \bar{v}_1}{t_k} &= \frac{\nabla_1 F(x_k)^* y_k - \nabla_1 F(\bar{x})^* \bar{y}}{t_k} \\ &= \left(\frac{\nabla_1 F(\bar{x} + t_k \xi_k)^* - \nabla_1 F(\bar{x})^*}{t_k} \right) y_k + \nabla_1 F(\bar{x})^* \left(\frac{y_k - \bar{y}}{t_k} \right), \end{aligned}$$

where the matrix quotient converges to a certain matrix A (depending only on $\bar{x}$ and ξ), inasmuch as F is C^2. This tells us that

$$\nabla_1 F(\bar{x})^* \left(\frac{y_k - \bar{y}}{t_k} \right) \to v_1' - d_1, \;\text{ where }\; d_1 = A\bar{y}.$$

Further, the analysis reveals that for *any* sequence of vectors $\widetilde{y}_k \to \bar{y}$ with

$$\widetilde{y}_k \in \partial g(F(x_k)) \;\text{ and }\; \lim_{k\to\infty} \nabla_1 F(\bar{x})^* \left(\frac{\widetilde{y}_k - \bar{y}}{t_k} \right) = v_1' - d_1$$

the vectors $\widetilde{v}_{1,k} := \nabla_1 F(x_k)^* \widetilde{y}_k$ converge to $\bar{v}_1$ while $(\widetilde{v}_{1,k} - \bar{v}_1)/t_k$ converges to v_1'.
Recall that the set $\operatorname{dom} g$ has a regular C^2 representation at $F(\bar{x})$. This means that there exist a neighborhood U of $F(\bar{x})$ along with C^2 functions h_j for $j = 1, \ldots, r$ and an integer $q \in [0, r]$ such that $\operatorname{dom} g \cap U$ is the set of points $u \in U$ such that

$$h_j(u) - h_j(F(\bar{x})) \begin{cases} \leq 0 & \text{for } j = 1, \ldots, q, \\ = 0 & \text{for } j = q+1, \ldots, r, \end{cases}$$

and this system satisfies the Mangasarian-Fromovitz constraint qualification at $F(\bar{x})$ i.e., that the only $\mu = (\mu_1, \dots, \mu_r)$ with $\mu_j \geq 0$ for $j = 1, \dots, q$ and $\sum_{j=1}^{r} \mu_j \nabla h_j(F(\bar{x})) = 0$ is $\mu = (0, \dots, 0)$.
It is well known, see for example [21, Thm. 4.3], that for $u \in \operatorname{dom} g \cap U$ we have

$$N_{\operatorname{dom} g}(u) = \left\{ \sum_{j=1}^{r} \mu_j \nabla h_j(u) \,\middle|\, \begin{array}{l} \mu_j \geq 0 \ \text{ for } \ j = 1, \dots, q \ \text{ with } \ h_j(u) = h_j(F(\bar{x})) \\ \mu_j = 0 \ \text{ for } \ j = 1, \dots, q \ \text{ with } \ h_j(u) < h_j(F(\bar{x})) \end{array} \right\}.$$

By Proposition 2.1, there exist for $i = 1, \dots, s$, $j = 1, \dots, r$, and each k, scalars $\lambda_{i,k} \in [0, 1]$ with $\sum_{i=1}^{s} \lambda_{i,k} = 1$ and $\mu_{j,k}$ with $\mu_{j,k} \geq 0$ for $j = 1, \dots, q$ such that

$$y_k = \sum_{i=1}^{s} \lambda_{i,k} \nabla g_i(F(x_k)) + \sum_{j=1}^{r} \mu_{j,k} \nabla h_j(F(x_k)). \tag{2.1}$$

We may assume without loss of generality that $\lambda_{i,k} \to \lambda_i$ (as $k \to \infty$), with $\lambda_i \in [0, 1]$, $\sum_{i=1}^{s} \lambda_i = 1$. We may also assume that $\mu_{j,k} \to \mu_j$ as $k \to \infty$. (If not we, could divide (2.1) by $\rho_k = \max\{|\mu_{j,k}|\}$. Obviously $\lambda_{i,k}/\rho_k \to 0$. In the limit we would therefore have a sum of multiples of the $\nabla h_j(F(\bar{x}))$'s equaling 0 and with the coefficient of $\nabla h_j(F(\bar{x}))$ for $1 \leq j \leq q$ nonnegative. But by the Mangasarian-Fromovitz constraint qualification, all coefficients must be 0, a contradiction.) Therefore

$$\bar{y} = \sum_{i=1}^{s} \lambda_i \nabla g_i(F(\bar{x})) + \sum_{j=1}^{r} \mu_j \nabla h_j(F(\bar{x})).$$

Under our assumptions, the vectors

$$\nabla_1 F(\bar{x})^* \left[\sum_{i=1}^{s} \lambda_{i,k} \left[\frac{\nabla g_i(F(x_k)) - \nabla g_i(F(\bar{x}))}{t_k} \right] + \sum_{j=1}^{r} \mu_{j,k} \left[\frac{\nabla h_j(F(x_k)) - \nabla h_j(F(\bar{x}))}{t_k} \right] \right]$$

converge to a certain vector d_1'. Note that d_1' only depends on $\bar{x}$, ξ, the λ_i's and the μ_j's. Consequently,

$$\begin{aligned} &\lim_{k \to \infty} \nabla_1 F(\bar{x})^* \left[\sum_{i=1}^{s} \left[\frac{\lambda_{i,k} - \lambda_i}{t_k} \right] \nabla g_i(F(\bar{x})) + \sum_{j=1}^{r} \left[\frac{\mu_{j,k} - \mu_j}{t_k} \right] \nabla h_j(F(\bar{x})) \right] \\ &= v_1' - d_1 - d_1'. \end{aligned} \tag{2.2}$$

Our goal is to replace $\lambda_{i,k}$ by $\widetilde{\lambda}_{i,k}$ with $\widetilde{\lambda}_{i,k}$ converging to λ_i, and $\mu_{j,k}$ by $\widetilde{\mu}_{j,k}$ with $\widetilde{\mu}_{j,k} \to \mu_j$. With

$$\widetilde{y}_k := \left[\sum_{i=1}^{s} \widetilde{\lambda}_{i,k} \nabla g_i(F(x_k)) + \sum_{j=1}^{r} \widetilde{\mu}_{j,k} \nabla h_j(F(x_k)) \right],$$

we will choose $\widetilde{\lambda}_{i,k}$ and $\widetilde{\mu}_{j,k}$ so that

$$\widetilde{y}_k \in \partial g(F(x_k)).$$

In addition we'll have as in (2.2)

$$\lim_{k\to\infty} \nabla_1 F(\bar{x})^* \left[\sum_{i=1}^{s}\left[\frac{\widetilde{\lambda}_{i,k}-\lambda_i}{t_k}\right]\nabla g_i(F(\bar{x})) + \sum_{j=1}^{r}\left[\frac{\widetilde{\mu}_{j,k}-\mu_j}{t_k}\right]\nabla h_j(F(\bar{x}))\right]$$
$$= v_1' - d_1 - d_1'.$$

Moreover we will show that

$$\sum_{i=1}^{s}\left[\frac{\widetilde{\lambda}_{i,k}-\lambda_i}{t_k}\right]\nabla g_i(F(\bar{x})) + \sum_{j=1}^{r}\left[\frac{\widetilde{\mu}_{j,k}-\mu_j}{t_k}\right]\nabla h_j(F(\bar{x}))$$

converges to some vector θ_1; we then have $\nabla_1 F(\bar{x})^*(\bar{x})\theta_1 = v_1' - d_1 - d_1'$. It follows that $\widetilde{y}_k \to \bar{y}$. Then with $\widetilde{v}_{j,k} = \nabla_j F(x_k)^*\widetilde{y}_k$ for $j = 1, 2$ we'll have (of course) $(\widetilde{v}_{1,k}-\bar{v}_1)/t_k \to v_1'$ and with $\bar{v}_2 := \nabla_2 F(\bar{x})^*\bar{y}$, it is absolutely elementary to show that $(\widetilde{v}_{2,k}-\bar{v}_2)/t_k$ converges to some v_2'. The proof will therefore be completed once we show that we can find $\widetilde{\lambda}_{i,k}$ and $\widetilde{\mu}_{j,k}$ with the above properties.

The proof now proceeds by induction on the number of indices with $(\lambda_{i,k}-\lambda_i)/t_k$ or $(\mu_{j,k}-\mu_j)/t_k$ unbounded (as $k\to\infty$). Let N be the number of such indices. Let

$$v_i = \nabla_1 F(\bar{x})^*\nabla g_i(F(\bar{x})), \quad d_j = \nabla_1 F(\bar{x})^*\nabla h_j(F(\bar{x})),$$

$$a_{i,k} = \frac{\lambda_{i,k}-\lambda_i}{t_k} \quad\text{and}\quad b_{j,k} = \frac{\mu_{j,k}-\mu_j}{t_k}.$$

Note that by our assumptions

$$\sum_{i=1}^{s} a_{i,k}v_i + \sum_{j=1}^{r} b_{j,k}d_j \tag{2.3}$$

is bounded as $k\to\infty$ (by (2.2) it actually converges to $v_1' - d_1 - d_1'$). Also note that

$$\sum_{i=1}^{s} a_{i,k} = 0, \tag{2.4}$$

because $\sum_{i=1}^{s}\lambda_{i,k} = 1 = \sum_{i=1}^{s}\lambda_i$.

N=1: From (2.4) we can not have only one $a_{i,k}$ unbounded (if this is the case assume without loss of generality that $a_{1,k}$ is unbounded. When we divide (2.4) by $a_{1,k}$, in the limit we have that $1 = 0$; hmm!). Since there is only a finite number of indices, me may assume by relabeling if necessary that $b_{1,k}$ is unbounded. From (2.3) we conclude that $d_1 = 0$. Therefore simply let $\widetilde{\mu}_{1,k} = \mu_1$.

Assume true $N = 1, 2, \ldots, m-1$. Let $N = m$. Let

$$\alpha_k = \max\left\{\max_{i=1,\ldots,s} |a_{i,k}|, \max_{j=1,\ldots,r} |b_{j,k}|\right\}.$$

Assume (without loss of generality) that $(a_{i,k}/\alpha_k) \to a_i$ and that $(b_{j,k}/\alpha_k) \to b_j$. Note that a_i equals 0 when $a_{i,k}$ is bounded and that b_j equals 0 when $b_{j,k}$ is bounded. From (2.3) and (2.4) we have

$$\sum_{i=1}^{s} a_i v_i + \sum_{j=1}^{r} b_j d_j = 0 \tag{2.5}$$

and

$$\sum_{i=1}^{s} a_i = 0. \tag{2.6}$$

Let

$$\beta_k^0 = \min\left\{1, \min\left\{\frac{\lambda_{i,k}}{t_k \alpha_k a_i} \,\middle|\, 1 \le i \le s, \lambda_i = 0, a_i \ne 0\right\}\right\}.$$

If there is i with $\lambda_i = 1$ and $a_i \ne 0$ then let

$$\beta_k^1 = \min\left\{\beta_k^0, \frac{\lambda_{i,k} - 1}{t_k \alpha_k a_i}\right\},$$

if not let $\beta_k^1 = \beta_k^0$. Let

$$\beta_k^2 = \min\left\{1, \min\left\{\frac{\mu_{j,k}}{t_k \alpha_k b_j} \,\middle|\, 1 \le j \le q, \mu_j = 0, b_j \ne 0\right\}\right\}.$$

Finally let

$$\beta_k = \min\{\beta_k^1, \beta_k^2\}.$$

Note that $\beta_k \to 1$. In addition note that if $0 < \lambda_i < 1$ for all $i = 1, \ldots, s$ with $a_i \ne 0$ and $\mu_j > 0$ for $j = 1, \ldots, q$ with $b_j \ne 0$ then $\beta_k^0 = \beta_k^1 = \beta_k^2 = \beta_k = 1$ (this will be important in Claim 2.8).

For $i = 1, \ldots, s$ let

$$\widetilde{\lambda}_{i,k} = \lambda_{i,k} - t_k \alpha_k a_i \beta_k,$$

and for $j = 1, \ldots, r$ let

$$\widetilde{\mu}_{j,k} = \mu_{j,k} - t_k \alpha_k b_j \beta_k,$$

We may assume without loss of generality that $\widetilde{\lambda}_{i,k} \to \lambda_i$ and $\widetilde{\mu}_{j,k} \to \mu_j$ (because there is a only a finite number of indices i and j, there exists either some i with $t_k\alpha_k = (\substack{+\\-})(\lambda_{i,k} - \lambda_i)$ for all k in an infinite subset of $\mathbb{N}$ or some j with $t_k \alpha_k = (\substack{+\\-})(\mu_{j,k} - \mu_j)$ for all k in an infinite subset of $\mathbb{N}$; in either case a subsequence of $t_k\alpha_k$ converges to 0).

The role of β_k is to ensure that

$$\left[\sum_{i=1}^{s} \widetilde{\lambda}_{i,k} \nabla g_i(F(x_k)) + \sum_{j=1}^{r} \widetilde{\mu}_{j,k} \nabla h_j(F(x_k))\right] \in \partial g(F(x_k)). \tag{2.7}$$

We first show:

Claim 2.4. *For all k large enough and for all i, $\widetilde{\lambda}_{i,k} \in [0,1]$ and if $\lambda_{i,k} = 0$ then $\widetilde{\lambda}_{i,k} = 0$.*

Proof of Claim 2.4. First note that if $\lambda_{i,k} = 0$ for all k large, then $\lambda_i = 0$, $a_i = 0$ and $\widetilde{\lambda}_{i,k} = 0$ for all k large.

If $a_i = 0$ then $\widetilde{\lambda}_{i,k} = \lambda_{i,k}$. So assume that $a_i \neq 0$. If $\lambda_i \in (0,1)$ then eventually $\widetilde{\lambda}_{i,k} \in (0,1)$ because $t_k \alpha_k \to 0$ and $\lambda_{i,k} \to \lambda_i$.

If $\lambda_i = 0$ then for large k, $a_{i,k} = (\lambda_{i,k} - 0)/t_k > 0$ (if $\lambda_{i,k} = 0$ for all k large, then $a_i = 0$ a contradiction) so that $\alpha_k a_i > 0$ (recall that $a_{i,k}/(\alpha_k a_i)$ converges to 1). Therefore by the choice of β_k,

$$\widetilde{\lambda}_{i,k} = \lambda_{i,k} - t_k \alpha_k a_i \beta_k \geq \lambda_{i,k} - t_k \alpha_k a_i \frac{\lambda_{i,k}}{t_k \alpha_k a_i} = 0.$$

Clearly in this case $\widetilde{\lambda}_{i,k} \leq 1$ because $\alpha_k a_i > 0$ so that $\widetilde{\lambda}_{i,k} \leq \lambda_{i,k}$.

If $\lambda_i = 1$ then for large k, $a_{i,k} = (\lambda_{i,k} - 1)/t_k < 0$ (if $\lambda_{i,k} = 1$ for all k large, then $a_i = 0$ a contradiction) so that $\alpha_k a_i < 0$. This shows that $\widetilde{\lambda}_{i,k} \geq \lambda_{i,k}$. On the other hand

$$\widetilde{\lambda}_{i,k} = \lambda_{i,k} - t_k \alpha_k a_i \beta_k \leq \lambda_{i,k} - t_k \alpha_k a_i \left(\frac{\lambda_{i,k} - 1}{t_k \alpha_k a_i}\right) = 1.$$

□

Claim 2.5. $\sum_{i=1}^{s} \widetilde{\lambda}_{i,k} = 1$

Proof of Claim 2.5.

$$\begin{aligned}
\sum_{i=1}^{s} \widetilde{\lambda}_{i,k} &= \sum_{i=1}^{s} (\lambda_{i,k} - t_k \alpha_k a_i \beta_k) \\
&= \sum_{i=1}^{s} \lambda_{i,k} - t_k \alpha_k \beta_k \sum_{i=1}^{s} a_i \\
&= 1 - t_k \alpha_k \beta_k (0) \quad \text{(from (2.6))} \\
&= 1.
\end{aligned}$$

□

Claim 2.6. *For all k large enough and for all $1 \leq j \leq q$, $\widetilde{\mu}_{j,k} \geq 0$ and if $\mu_{j,k} = 0$ then $\widetilde{\mu}_{j,k} = 0$.*

Proof of Claim 2.6. First note that if $\mu_{j,k} = 0$ for all k large, then $\mu_j = 0$, $b_j = 0$ and $\widetilde{\mu}_{j,k} = 0$ for large k.
If $b_j = 0$ then $\widetilde{\mu}_{j,k} = \mu_{j,k} \geq 0$. So assume that $b_j \neq 0$. If $\mu_j > 0$ then eventually $\widetilde{\mu}_{j,k} > 0$ because $t_k\alpha_k \to 0$ and $\mu_{j,k} \to \mu_j$.
If $\mu_j = 0$ then $b_{j,k} = (\mu_{j,k} - 0)/t_k > 0$ (if $\mu_{j,k} = 0$ for all k large, then $b_j = 0$ a contradiction) so that $\alpha_k b_j > 0$. Therefore

$$\widetilde{\mu}_{j,k} = \mu_{j,k} - t_k\alpha_k b_j\beta_k \geq \mu_{j,k} - t_k\alpha_k b_j \frac{\mu_{j,k}}{t_k\alpha_k b_j} = 0.$$

□

By our subgradient assumptions on g, the combination of Claims 2.4, 2.5 and 2.6 shows that (2.7) is valid. Indeed by our choice of $\widetilde{\lambda}_{i,k}$ we have

$$\sum_{i=1}^{s} \widetilde{\lambda}_{i,k} \nabla g_i(F(x_k)) \in \mathrm{co}\{\nabla g_i(F(x_k)) \,|\, F(x_k) \in D_i\}$$

because $\sum_{i=1}^{s} \widetilde{\lambda}_{i,k} = 1$, $\widetilde{\lambda}_{i,k} \in [0,1]$ and $\widetilde{\lambda}_{i,k} = 0$ precisely when $\lambda_{i,k} = 0$ in other words when potentially $F(x_k) \notin D_i$. On the other hand

$$\sum_{j=1}^{r} \widetilde{\mu}_{j,k} \nabla h_j(F(x_k)) \in N_{\mathrm{dom}\, g}(F(x_k)),$$

because $\widetilde{\mu}_{j,k} \geq 0$ for $1 \leq j \leq q$ and $\widetilde{\mu}_{j,k} = 0$ precisely when $\mu_{j,k} = 0$ i.e. when potentially $h_j(F(x_k)) - h_j(F(\bar{x})) < 0$.

Claim 2.7.

$$\sum_{i=1}^{s} \Big(\frac{\widetilde{\lambda}_{i,k} - \lambda_i}{t_k}\Big) v_i + \sum_{j=1}^{r} \Big(\frac{\widetilde{\mu}_{j,k} - \mu_j}{t_k}\Big) d_j = \sum_{i=1}^{s} \Big(\frac{\lambda_{i,k} - \lambda_i}{t_k}\Big) v_i + \sum_{j=1}^{r} \Big(\frac{\mu_{j,k} - \mu_j}{t_k}\Big) d_j.$$

Proof of Claim 2.7.

$$\begin{aligned}
\sum_{i=1}^{s} (\frac{\widetilde{\lambda}_{i,k} - \lambda_i}{t_k}) v_i + \sum_{j=1}^{r} \Big(\frac{\widetilde{\mu}_{j,k} - \mu_j}{t_k}\Big) d_j &= \sum_{i=1}^{s} \Big(\frac{\lambda_{i,k} - \lambda_i}{t_k}\Big) v_i + \sum_{j=1}^{r} \Big(\frac{\mu_{j,k} - \mu_j}{t_k}\Big) d_j \\
&\quad - \beta_k\alpha_k \left[\sum_{i=1}^{s} a_i v_i + \sum_{j=1}^{r} b_j d_j\right] \\
&= \sum_{i=1}^{s} \Big(\frac{\lambda_{i,k} - \lambda_i}{t_k}\Big) v_i + \sum_{j=1}^{r} \Big(\frac{\mu_{j,k} - \mu_j}{t_k}\Big) d_j \\
&\qquad \text{(cf. (2.5))}
\end{aligned}$$

□

Claim 2.8. *Without any loss of generality, we may assume that there exists some i with $\widetilde{\lambda}_{i,k} = \lambda_i$ for all k or some j with $\widetilde{\mu}_{j,k} = \mu_j$ for all k.*

Proof of Claim 2.8. Fix $k \in \mathbb{N}$. If $\beta_k = \dfrac{\lambda_{i,k}}{t_k \alpha_k a_i}$ for some i with $\lambda_i = 0$ and $a_i \neq 0$, then for such i,

$$\widetilde{\lambda}_{i,k} = \lambda_{i,k} - t_k \alpha_k a_i \beta_k = 0 = \lambda_i.$$

If $\beta_k = \dfrac{\mu_{j,k}}{t_k \alpha_k b_j}$ for some j with $\mu_j = 0$, $1 \leq j \leq q$, and $b_j \neq 0$, then for such j,

$$\widetilde{\mu}_{j,k} = \mu_{j,k} - t_k \alpha_k b_j \beta_k = 0 = \mu_j.$$

If $\beta_k = \dfrac{\lambda_{i,k} - 1}{t_k \alpha_k a_i}$ for some i with $\lambda_i = 1$ and $a_i \neq 0$, then

$$\widetilde{\lambda}_{i,k} = \lambda_{i,k} - t_k \alpha_k a_i \left(\frac{\lambda_{i,k} - 1}{t_k \alpha_k a_i} \right) = 1 = \lambda_i.$$

If any of the three cases mentioned above is true for all k in an infinite subset of $\mathbb{N}$, then we are done because there is only a finite number of indices i and j (some index i or j would be repeated infinitely often, then take an appropriate subsequence). If not this means that $\beta_k = 1$ for all k sufficiently large (see the comment made after the introduction of the β_k's). Again because there is only a finite number of indices we may assume by taking a subsequence if necessary that there exists some i with $\alpha_k = |a_{i,k}|$ for all k or there exists some j with $\alpha_k = |b_{j,k}|$ for all k. If $\alpha_k = |a_{i,k}|$ then $a_i \alpha_k = a_{i,k}$ (because $a_{i,k}/\alpha_k = a_{i,k}/|a_{i,k}| \to a_i$), and

$$\begin{aligned}
\widetilde{\lambda}_{i,k} &= \lambda_{i,k} - t_k \alpha_k a_i \\
&= \lambda_{i,k} - t_k a_{i,k} \\
&= \lambda_{i,k} - (\lambda_{i,k} - \lambda_i) \\
&= \lambda_i,
\end{aligned}$$

for all k. If on the other hand $\alpha_k = |b_{j,k}|$, then by a similar argument $\widetilde{\mu}_{j,k} = \mu_j$ for all k. □

So finally we have produced $\widetilde{\lambda}_{i,k}$ and $\widetilde{\mu}_{j,k}$ with $\widetilde{\lambda}_{i,k} \to \lambda_i$, $\widetilde{\mu}_{j,k} \to \mu_j$ and

$$\widetilde{y}_k := \left[\sum_{i=1}^{s} \widetilde{\lambda}_{i,k} \nabla g_i(F(x_k)) + \sum_{j=1}^{r} \widetilde{\mu}_{j,k} \nabla h_j(F(x_k)) \right] \in \partial g(F(x_k)).$$

By Claim 2.7, for all k

$$\nabla_1 F(\bar{x})^* \left(\frac{\widetilde{y}_k - \bar{y}}{t_k} \right) = \nabla_1 F(\bar{x})^* \left(\frac{y_k - \bar{y}}{t_k} \right).$$

And the number of indices with $\dfrac{\widetilde{\lambda}_{i,k} - \lambda_i}{t_k}$ or $\dfrac{\widetilde{\mu}_{j,k} - \mu_j}{t_k}$ unbounded is strictly less than m (Claim 2.8). This completes the proof. $\square$

References

[1] R. T. Rockafellar: Proto-differentiability of set-valued mappings and its applications in optimization, in: Analyse Non Linéaire, H. Attouch et al. (eds.), Gauthier-Villars, Paris, 1989, 449–482.

[2] R. T. Rockafellar: Generalized second derivatives of convex functions and saddle functions, Trans. Amer. Math. Soc. 320 (1990) 810–822.

[3] R. T. Rockafellar: Nonsmooth analysis and parametric optimization, in: Methods of Nonconvex Analysis (A. Cellina, ed.), Springer-Verlag Lecture Notes in Math. 1446 (1990), 137–151.

[4] R. A. Poliquin: Proto-differentiation of subgradient set-valued mappings, Canadian J. Math. 42 (1990) 520–532.

[5] R. Cominetti: On pseudo-differentiability, Trans. Amer. Math. Soc. 324 (1991) 843–865.

[6] R. A. Poliquin: An extension of Attouch's theorem and its application to second-order epi-differentiation of convexly composite functions, Trans. Amer. Math. Soc. 332 (1992) 861–874.

[7] R. A. Poliquin, R. T. Rockafellar: Amenable functions in optimization, in: Nonsmooth Optimization Methods and Applications (F. Giannessi, ed.), Gordon and Breach, Philadelphia, 1992, 338–353.

[8] R. A. Poliquin, R. T. Rockafellar: A calculus of epi-derivatives applicable to optimization, Canadian J. Math. 45 (1993) 879–896.

[9] R. A. Poliquin, R. T. Rockafellar: Proto-derivative formulas for basic subgradient mappings in mathematical programming, Set-Valued Analysis 2 (1994) 275–290.

[10] A. Levy, R. A. Poliquin, L. Thibault: Partial extension of Attouch's theorem with applications to proto-derivatives of subgradient mappings, Trans. Amer. Math. Soc. 347 (1995) 1269–1294.

[11] R. A. Poliquin, R. T. Rockafellar: Second-order nonsmooth analysis in nonlinear programming, in: Recent Advances in Nonsmooth Optimization (D. Du, L. Qi, and R. Womersley, eds.), World Scientific Publishers, 1995, 322–350.

[12] R. A. Poliquin, R. T. Rockafellar: Prox-regular functions in variational analysis, Trans. Amer. Math. Soc. 348 (1996) 1805–1838.

[13] R. A. Poliquin, R. T. Rockafellar: Generalized Hessian properties of regularized nonsmooth functions, SIAM J. Optimization 6 (1996) 1121–1137.

[14] A. B. Levy, R. T. Rockafellar: Sensitivity analysis of solutions to generalized equations, Trans. Amer. Math. Soc. 345 (1994) 661–671.

[15] R. T. Rockafellar: First- and second-order epi-differentiability in nonlinear programming, Trans. Amer. Math. Soc. 307 (1988) 75–107.

[16] R. T. Rockafellar, R. J-B Wets: Variational Analysis, Springer-Verlag, Berlin et al., 1997.

[17] A. B. Levy, R. T. Rockafellar: Variational conditions and the proto-differentiation of partial subgradient mappings, Nonlinear Anal. Th. Meth. Appl. 26 (1996) 1951–1964.

[18] J. Sun: On Monotropic Piecewise Quadratic Programming, Ph.D. Dissertation, University of Washington, Seattle, 1986.

[19] D. Azé, R. A. Poliquin: Equi-calmness and epi-derivatives that are pointwise limits, J. Math. Anal. Appl., to appear.

[20] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton, NJ, 1970.

[21] R. T. Rockafellar: Lagrange multipliers and optimality, SIAM Review 35 (1993) 183–238.