PROTO-DERIVATIVE FORMULAS FOR BASIC SUBGRADIENT MAPPINGS IN MATHEMATICAL PROGRAMMING

April, 1993

R. A. Poliquin and R. T. Rockafellar*

Abstract. Subgradient mappings associated with various convex and nonconvex functions are a vehicle for stating optimality conditions, and their proto-differentiability plays a role therefore in the sensitivity analysis of solutions to problems of optimization. Examples of special interest are the subgradients of the max of finitely many $\mathcal{C}^2$ functions, and the subgradients of the indicator of a set defined by finitely many $\mathcal{C}^2$ constraints satisfying a basic constraint qualification. In both cases the function has a property called full amenability, so the general theory of existence and calculus of proto-derivatives of subgradient mappings associated with fully amenable functions is applicable. This paper works out the details for such examples. A formula of Auslender and Cominetti in the case of a max function is improved in particular.

Keywords. Proto-derivatives, generalized second derivatives, nonsmooth analysis, epi-derivatives, subgradient mappings, amenable functions.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 49A52, 58C06, 58C20; Secondary 90C30

---

* 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–9200303 for the second author.

1. Introduction

A set-valued mapping $\Gamma : I\!\!R^n \rightrightarrows I\!\!R^m$ is *proto-differentiable* [1] at a point x and for a particular element $v \in \Gamma(x)$ if the set-valued mappings

$$\Delta_{x,v,t} : \xi \mapsto \big[\Gamma(x + t\xi) - v\big]/t,$$

regarded as a family indexed by $t > 0$, graph-converge as $t \searrow 0$ (i.e., set convergence of the graphs; see section 3). If so, the limit mapping is denoted by $\Gamma'_{x,v}$ and called the *proto-derivative* of Γ at x for v. It assigns to each $\xi \in I\!\!R^n$ a subset $\Gamma'_{x,v}(\xi)$ of $I\!\!R^n$, which could be empty for some choices of ξ.

A key issue in parametric optimization is the proto-differentiability of the mapping that associates with each vector of parameters the corresponding set of optimal solutions, or in nonconvex programming perhaps some set of "quasi-optimal solutions" expressed by a system of conditions related to optimality. Typical examples involve first-order optimality conditions in terms of the subgradients of the essential objective function in a given problem. The question then comes down to whether the subgradient mapping of such an objective function is proto-differentiable. These motivations in sensitivity analysis are explained in [2], [3], and [4].

Many subgradient mappings are known to be proto-differentiable. In fact the subgradient mapping ∂f of any fully amenable function f is proto-differentiable, as proved by Poliquin [5]. A function $f : I\!\!R^n \to \overline{I\!\!R}$ is *amenable* at $\bar{x}$, a point where $f(\bar{x})$ is finite, if on some open neighborhood V of $\bar{x}$ there is a $\mathcal{C}^2$ mapping $F : V \to I\!\!R^m$ and a convex, lower semicontinuous function $g : I\!\!R^m \to \overline{I\!\!R}$ such that $f(x) = g\big(F(x)\big)$ for $x \in V$ and the following condition (an abstract constraint qualification) is satisfied at $\bar{x}$:

$$\text{there is no vector } y \neq 0 \text{ in } N_{\operatorname{dom} g}\big(F(\bar{x})\big) \text{ with } \nabla F(\bar{x})^* y = 0. \tag{1.1}$$

Here $\nabla F(\bar{x})$ denotes the $m \times n$ Jacobian matrix of F at $\bar{x}$, and $\nabla F(\bar{x})^*$ is its transpose. Further, $N_{\operatorname{dom} g}\big(F(\bar{x})\big)$ denotes the normal cone to the nonempty convex set $\operatorname{dom} g$ at the point $F(\bar{x})$. (When $F(\bar{x}) \in \operatorname{int}(\operatorname{dom} g)$ this cone consists just of the vector 0, and condition (1.1) is then satisfied trivially.)

For f to be *fully* amenable, the assumption is added that g can be chosen to be *piecewise linear-quadratic*. This means that $\operatorname{dom} g$ (the set of points where the value of g is not ∞) can be expressed as the union of a finite collection of polyhedral (convex) sets, on each of which g is given by a polynomial expression with no terms higher than degree two. For more on amenable and fully amenable functions, see [2], [3], [5]–[8].

Examples of piecewise linear-quadratic convex functions g are *polyhedral* functions-—having polyhedral epigraph—such as the indicator function and support function of a polyhedral set, or the max of a finite collection of affine functions. Such functions are merely piecewise linear. On the other hand, a function giving the Euclidean distance squared from a polyhedral set is piecewise linear-quadratic but not polyhedral.

Full amenability is a local property, in the sense that when it holds at $\bar{x}$ it actually holds at all points x in some neighborhood of $\bar{x}$ relative to $\operatorname{dom} f$. In speaking of subgradients $v \in \partial f(x)$ of a fully amenable function f, which need not be convex, we are able to take advantage of the fact that such functions are Clarke regular. For Clarke regular functions, the various definitions of $\partial f(x)$ (cf. Clarke [9], Mordukhovich [10] and Rockafellar [11] in particular) all agree.

The class of fully amenable functions is much larger than might at first be apparent. Many examples of importance in mathematical programming have been indicated in [2], [3], [6]–[8]. In this note we focus on two that are central: the pointwise maximum of a collection of finitely many C^2 functions and the indicator of a set defined by finitely many C^2 constraints under a constraint qualification. We also look at the essential objective function of a smooth nonlinear programming problem having such a system of constraints.

Example 1. *Let f be specified by*

$$f(x) = \max\{\, f_1(x), \ldots, f_m(x)\,\}, \tag{1.2}$$

where each function $f_i : \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{C}^2$. Then f is everywhere fully amenable.

To see this, simply observe that $f(x) = g\big(F(x)\big)$ for

$$F(x) = \big(f_1(x), \ldots, f_m(x)\big), \qquad g(w_1, \ldots, w_m) = \max\{w_1, \ldots, w_m\}, \tag{1.3}$$

and note that g is polyhedral. Condition (1.1) is automatically satisfied, since $\operatorname{dom} g$ is all of $\mathbb{R}^m$ and therefore contains every point $F(x)$ in its interior.

Example 2. *Let f be an indicator function of the form*

$$\begin{gathered} f(x) = \delta_C(x) := \begin{cases} 0 & \text{if } x \in C, \\ \infty & \text{if } x \notin C, \text{ where} \end{cases} \\ C := \big\{ x \in X \,\big|\, f_i(x) \in I_i,\ i = 1 \ldots m \big\}, \end{gathered} \tag{1.4}$$

under the assumption that X is a polyhedral set in $\mathbb{R}^n$, and for each i, $f_i : \mathbb{R}^n \to \mathbb{R}$ is a function of class C^2 while I_i is a closed interval in $\mathbb{R}$. Then f is fully amenable at any point $\bar{x} \in C$ at which the following basic constraint qualification is satisfied:

$$\begin{cases} \text{the only multipliers } y_i \in N_{I_i}\big(f_i(\bar{x})\big) \text{ satisfying} \\ -\sum_{i=1}^m y_i \nabla f_i(\bar{x}) \in N_X(\bar{x}) \text{ are } y_1 = 0, \ldots, y_m = 0. \end{cases} \tag{1.5}$$

The composite representation $f = g \circ F$ that yields the conclusion of full amenability in Example 2 has

$$\begin{aligned} F(x) &= \big(f_1(x), \ldots, f_m(x), x\big), \\ g &= \delta_D \ \text{ for } \ D := I_1 \times \cdots \times I_m \times X. \end{aligned} \tag{1.6}$$

Then $N_D(F(\bar{x})) = N_{I_1}\big(f_1(\bar{x})\big) \times \cdots \times N_{I_m}\big(f_m(\bar{x})\big) \times N_X\big(\bar{x}\big)$. Since D is a polyhedral set, g is once more a polyhedral function. Again it should be noted that if the constraint qualification holds at $\bar{x}$, it must actually hold at all points of C in some neighborhood of $\bar{x}$. When f is an indicator δ_C, the subgradient set $\partial f(x)$ is the normal cone $N_C(x)$ to C at x (replaced by the empty set when $x \notin C$).

Insight into the constraint qualification (1.5) is gained from the classical case where X is the whole space, $I_i = (-\infty, 0]$ for $i = 1, \ldots, s$, (so that $N_{I_i}\big(f_i(\bar{x})\big)$ equals $[0, \infty)$ if $f_i(\bar{x}) = 0$ but equals $\{0\}$ if $f_i(\bar{x}) < 0$), and $I_i = \{0\}$ for $i = s+1, \ldots, m$ (so that for such indices $N_{I_i}\big(f_i(x)\big) = (-\infty, \infty)$ as long as $f_i(x) = 0$). The condition is then the dual statement of the familiar Mangasarian-Fromovitz constraint qualification. More generally, when I_i is the closed interval with lower bound a_i and upper bound b_i (these bounds possibly being infinite), with $a_i < b_i$, the relation $y_i \in N_{I_i}\big(f_i(x)\big)$ specifies the sign of the multiplier y_i in the following pattern, depending on whether the constraint $f_i(x) \in [a_i, b_i]$ is satisfied with $f_i(x)$ at either bound or in between:

$$y_i \in N_{[a_i,b_i]}\big(f_i(x)\big) \iff \begin{cases} y_i \geq 0 & \text{when } a_i < f_i(x) = b_i, \\ y_i \leq 0 & \text{when } a_i = f_i(x) < b_i, \\ y_i = 0 & \text{when } a_i < f_i(x) < b_i. \end{cases} \tag{1.7}$$

The following variant of Example 2 adds a $\mathcal{C}^2$ objective function to the indicator δ_C.

Example 3. *Suppose*

$$f(x) = f_0(x) + \delta_C(x) = \begin{cases} f_0(x) & \text{if } x \in C, \\ \infty & \text{if } x \notin C, \end{cases}$$

where f_0 is of class $\mathcal{C}^2$ and the set C has the form in Example 2. Then f is fully amenable at any point $\bar{x} \in C$ at which the constraint qualification (1.5) is satisfied.

This time the composite representation $f = g \circ F$ to take in verifying the asserted full amenability is

$$\begin{aligned} &F(x) = \big(f_0(x), f_1(x), \ldots, f_m(x), x\big), \\ &g(u_0, u_1, \ldots, u_m, x) = u_0 + \delta_D\big(u_1, \ldots, u_m, x\big), \end{aligned} \tag{1.8}$$

for the same set D as in Example 2. Here too g is polyhedral.

Proto-derivatives of ∂f for the max functions f in Example 1 have been studied by Auslender and Cominetti [12] and Penot [13]. Auslender and Cominetti obtained a formula for the mapping that corresponds to the "outer graphical limit" of the difference quotient mappings in the definition of proto-differentiability, but they did not show that the outer and corresponding "inner" limits coincide and thus did not establish proto-differentiability itself. In Penot's work the setting is potentially infinite-dimensional, and proto-differentiability is proved only under a sharp restriction. None of these authors utilized the composite representation (1.3), as we do here. Anyway, the formula we obtain here for Example 1 is simpler than theirs and does not require any extra assumptions. Expressions for the proto-derivatives of the subgradient mappings in Examples 2 and 3 have not previously been developed.

A relationship of fundamental importance in determining the proto-derivatives of the mapping $\partial f : I\!R^n \rightrightarrows I\!R^n$, when f is fully amenable, arises from a theory of generalized second derivatives developed in Rockafellar [6], [7], [14], [15]. This theory utilizes epi-convergence instead of pointwise convergence of second-order difference quotients, where epi-convergence of a sequence of functions refers to set convergence of their epigraphs.

A lower semicontinuous function $f : I\!R^n \to \overline{I\!R}$ is said to be *epi-differentiable*, at a point x where $f(x)$ is finite, if the first-order difference quotient functions $\Delta_{x,t} f : I\!R^n \to \overline{I\!R}$ defined by

$$\Delta_{x,t} f(\xi) = \big[f(x + t\xi) - f(x)\big]/t \ \text{ for } \ t > 0$$

epi-converge as $t \searrow 0$, the limit being a proper function (somewhere finite, nowhere $-\infty$). This limit is then the *epi-derivative* function f'_x. In like manner, f is *twice epi-differentiable* at x for a vector $v \in I\!R^n$ if it is epi-differentiable at x and the second-order difference quotient functions $\Delta^2_{x,v,t} f : I\!R^n \to \overline{I\!R}$ defined by

$$\Delta^2_{x,v,t} f(\xi) = \big[f(x + t\xi) - f(x) - t\langle v, x\rangle\big]/\tfrac{1}{2}t^2 \ \text{ for } \ t > 0$$

epi-converge to a proper function as $t \searrow 0$. The limit is then the *second epi-derivative* function $f''_{x,v}(\xi)$.

Rockafellar established in [6] that when f is fully amenable at x, it is twice epi-differentiable at x for every $v \in \partial f(x)$. On the other hand, he showed in [15] that a general convex function f is twice epi-differentiable at x for a vector $v \in \partial f(x)$ if and only if ∂f is proto-differentiable at x for v, and then

$$(\partial f)'_{x,v}(\xi) = \partial\big(\tfrac{1}{2} f''_{x,v}\big)(\xi) \ \text{ for all } \ \xi. \tag{1.9}$$

(For an infinite-dimensional generalization see Do [16].) From these facts it follows that ∂f is proto-differentiable for any function f that is both fully amenable and convex. Poliquin [5] proved, however, that convexity is superfluous: for any fully amenable function f, the proto-derivatives of ∂f exist and can be determined through (1.9) from the formulas known for second-order epi-derivatives of f. (Again there is no need to distinguish between different definitions of subgradients in formula (1.9).) Proceeding on this basis, Rockafellar and Poliquin recently developed general calculus rules in [8] for the proto-derivatives of subgradient mappings, but the particular mappings associated with Examples 1, 2 and 3 were not explicitly treated in that work.

2. Specialized Formulas

The convex functions g in the representations $f = g \circ F$ underlying Examples 1, 2 and 3 are not just piecewise linear-quadratic but piecewise linear, i.e., polyhedral. In this case substantial simplifications are possible in formulas for the second-order epi-derivatives of f and proto-derivatives of ∂f. We extract from [5], [6] and [8] the facts that will be needed. For this purpose we denote by $y{\cdot}F$, for any vector $y \in I\!R^m$, the real-valued function defined on $I\!R^n$ by

$$(y{\cdot}F)(x) := \big\langle y, F(x) \big\rangle.$$

Theorem 1. *Suppose f is fully amenable at $\bar{x}$ and the function g in the local $f(x) = g\big(F(x)\big)$ composite representation in the definition can be taken to be polyhedral. Then for all x in a neighborhood of $\bar{x}$ relative to $\operatorname{dom} f$, f is Clarke regular at x with its subgradients given by*

$$\partial f(x) = \nabla F(x)^* \partial g\big(F(x)\big) = \big\{\, v \,\big|\, \exists\, y \in \partial g\big(F(x)\big) \text{ with } \nabla F(x)^* y = v \,\big\} \tag{2.1}$$

and first-order epi-derivatives given by

$$f'_x(\xi) = g'_{F(x)}\big(\nabla F(x)\xi\big). \tag{2.2}$$

For any $v \in \partial f(x)$ the second-order epi-derivatives of f at x for v exist and are given by

$$f''_{x,v}(\xi) = \max_{y \in Y(x,v)} \big\langle \xi, \nabla^2 (y{\cdot}F)(x)\xi \big\rangle + \delta_{\Xi(x,v)}(\xi), \tag{2.3}$$

where $Y(x,v)$ is a bounded, polyhedral set and $\Xi(x,v)$ is a polyhedral cone, namely

$$\begin{aligned} &Y(x,v) = \big\{\, y \in \partial g\big(F(x)\big) \,\big|\, \nabla F(x)^* y = v \,\big\}, \\ &\Xi(x,v) = N_{\partial f(x)}(v) = \big\{\, \xi \,\big|\, f'_x(\xi) = \langle v, \xi \rangle \,\big\} = \big\{\, \xi \,\big|\, f'_x(\xi) \le \langle v, \xi \rangle \,\big\}. \end{aligned} \tag{2.4}$$

Furthermore, ∂f is proto-differentiable at x for v with

$$(\partial f)'_{x,v}(\xi) = \begin{cases} \big\{\, \nabla^2 (y{\cdot}F)(x)\xi \,\big|\, y \in Y_{\max}(x,v,\xi) \,\big\} + N_{\Xi(x,v)}(\xi) & \text{if } \xi \in \Xi(x,v), \\ \emptyset & \text{if } \xi \notin \Xi(x,v), \end{cases} \tag{2.5}$$

where $Y_{\max}(x,v,\xi)$ is the closed face of $Y(x,v)$ consisting of the multiplier vectors y that achieve the maximum in (2.3).

Proof. Only the special implications of the polyhedral nature of g need to be addressed, since everything else is in the references cited; see [8, Sec. 2] for a synopsis. In general, the term

$$\gamma_{F(x)}\big(\nabla F(x)\xi\big) = \lim_{t \searrow 0} \Big[g\Big(F(x) + t\nabla F(x)\xi\Big) - g\Big(F(x)\Big) - t\langle v, \xi\rangle\Big] \Big/ \tfrac{1}{2}t^2$$

would have to be added to the formula in (2.3) according to [6, Theorem 4.5]. But this vanishes when g is polyhedral and consequently piecewise linear relative to the polyhedral set $\operatorname{dom} g$. Thus, (2.3) is correct.

Equation (2.5) can be derived from (1.9) by ordinary subdifferential calculus as applied to (2.3). Denote the quadratic function of ξ in the max expression in (2.3) by $Q_y(\xi)$, observing that this depends linearly on y. In this notation $f''_{x,v}(\xi) = h(\xi) + \delta_{\Xi(x,v)}(\xi)$ for $h = \max_{y \in Y(x,v)} Q_y$. Therefore $\partial f''_{x,v}(\xi) = \partial h(\xi) + N_{\Xi(x,v)}(\xi)$, where

$$\partial h(\xi) = \Big\{ \nabla Q_y(\xi) \,\Big|\, y \in Y_{\max}(x,v,\xi) \Big\} \text{ with } \nabla Q_y(\xi) = 2\nabla^2 (y{\cdot}F)(x)\xi.$$

Invoking (1.9), we conclude that (2.5) holds. □

Corollary 1. *Under the assumptions in Theorem 1 each of the following properties implies all the others:*

(a) $(\partial f)'_{x,v}(\xi)$ *is nonempty for every* $\xi \in I\!R^n$,

(b) $(\partial f)'_{x,v}(\xi)$ *is bounded for every* $\xi \in I\!R^n$,

(c) $f''_{x,v}(\xi)$ *is finite for every* $\xi \in I\!R^n$,

(d) $\partial f(x) = \{v\}$,

(e) f *is differentiable at* x *with* $\nabla f(x) = v$.

Proof. As seen from the formulas in the theorem, all these properties are equivalent to having $\Xi(x,v) = I\!R^n$. □

We are ready now to treat Examples 1, 2 and 3 one by one. In every case the first-order results are well known, but we include them in the theorem statement as an aid to clarifying the context and fixing the notation.

Theorem 2. *In Example 1, consider any $x \in I\!R^n$ and let $I(x)$ denote the set of indices i such that $f_i(x) = f(x)$. Then*

$$\partial f(x) = \operatorname{co}\big\{ \nabla f_i(x) \,\big|\, i \in I(x) \big\}, \qquad f'_x(\xi) = \max_{i \in I(x)} \big\langle \nabla f_i(x), \xi \big\rangle. \tag{2.6}$$

For any $v \in \partial f(x)$ the second-order epi-derivatives of f at x for v exist and are given by

$$f''_{x,v}(\xi) = \max_{y \in Y(x,v)} \textstyle\sum_{i=1}^m y_i \big\langle \xi, \nabla^2 f_i(x)\xi \big\rangle \,+\, \delta_{\Xi(x,v)}(\xi), \tag{2.7}$$

where $Y(x,v)$ is a polyhedral set and $\Xi(x,v)$ is a polyhedral cone, namely

$$\begin{aligned} Y(x,v) = \Big\{ y \,\Big|\, & y_i \geq 0 \text{ if } i \in I(x),\ y_i = 0 \text{ if } i \notin I(x), \\ & \textstyle\sum_{i=1}^m y_i = 1,\ \sum_{i=1}^m y_i \nabla f_i(x) = v \Big\}, \\ \Xi(x,v) = \Big\{ \xi \,\Big|\, & \big\langle \nabla f_i(x) - v, \xi \big\rangle \leq 0 \text{ for all } i \in I(x) \Big\}. \end{aligned} \tag{2.8}$$

Furthermore, ∂f is proto-differentiable at x for v with

$$(\partial f)'_{x,v}(\xi) = \begin{cases} \Big\{ \sum_{i=1}^m y_i \nabla^2 f_i(x)\xi \,\Big|\, y \in Y_{\max}(x,v,\xi) \Big\} + K_{x,v}(\xi) & \text{if } \xi \in \Xi(x,v), \\ \emptyset & \text{if } \xi \notin \Xi(x,v), \end{cases} \tag{2.9}$$

where $Y_{\max}(x,v,\xi)$ is the closed face of $Y(x,v)$ consisting of the multiplier vectors y that achieve the maximum in (2.7), and $K_{x,v}(\xi)$ is the convex cone generated by the vectors $\nabla f_i(x) - v$ for $i \in I(x,\xi)$, this being the set of indices achieving the maximum in (2.6).

Proof. We are applying Theorem 1 to the case of F and g in (1.3), where

$$\begin{aligned} &\partial g(w) = \Big\{ y \,\Big|\, y_i \geq 0 \text{ if } w_i = g(w),\ y_i = 0 \text{ if } w_i < g(w),\ \textstyle\sum_{i=1}^m y_i = 1 \Big\}, \\ &g'_w(\omega) = \max\big\{ \omega_i \,\big|\, i \text{ such that } w_i = g(w) \big\}, \end{aligned}$$

along with

$$(y \cdot F)(x) = \sum\nolimits_{i=1}^m y_i f_i(x), \qquad \nabla F(x)^* y = \sum\nolimits_{i=1}^m y_i \nabla f_i(x). \tag{2.10}$$

Through these specializations the formulas in Theorem 1 turn into to the ones here, but in the case of $\Xi(x,v)$ this may not be obvious; we must verify also that $N_{\Xi(x,v)}(\xi)$ is the cone described as $K_{x,v}(\xi)$. From the definition of $\Xi(x,v)$ in (2.4) and the formula for $f'_x(\xi)$ in (2.6) we have

$$\begin{aligned} \xi \in \Xi(x,v) \quad &\Longleftrightarrow \quad \max_{i \in I(x)} \big\langle \nabla f_i(x), \xi \big\rangle \leq \langle v, \xi \rangle \\ &\Longleftrightarrow \quad \big\langle \nabla f_i(x) - v, \xi \big\rangle \leq 0 \text{ for all } i \in I(x), \end{aligned}$$

as claimed. Since $\Xi(x,v)$ is given this way by a system of linear constraints, its normal cone $N_{\Xi(x,v)}(\xi)$ at any of its elements ξ is the convex cone generated by the gradients of the constraints $\big\langle \nabla f_i(x) - v, \xi \big\rangle \leq 0$ that are active at ξ. Thus, this normal cone is the convex cone generated by the vectors $\nabla f_i(x) - v$ corresponding to the indices $i \in I(x)$ such that $\langle \nabla f_i(x) - v, \xi \rangle = 0$. In other words, it is $K_{x,v}(\xi)$. $\square$

The relationship between formula (2.9) and the results of Auslender and Cominetti [12] and Penot [13] will be discussed in Section 3.

Moving on now to Example 2, we denote by $T_C(x)$ the *tangent cone* to C at a point $x \in C$, and similarly by $T_X(x)$ the tangent cone to the polyhedral set X at x. These tangent cones are polar to the normal cones $N_C(x)$ and $N_X(x)$ (because we are dealing with convex sets or more generally sets that are Clarke regular, for which the various definitions in use for tangent cones all agree). The tangent cone notation will be useful also in handling constraints: we denote by $T_{I_i}(u_i)$ the tangent cone to the closed interval $I_i \subset I\!R$ at $u_i \in I_i$, which simply indicates the directions in which one can move from u_i without leaving I_i. Specifically, in parallel with (1.7), in the case where I_i has lower bound a_i and upper bound b_i (these possibly being infinite) with $a_i < b_i$, one has

$$T_{[a_i,b_i]}\big(f_i(x)\big) = \begin{cases} (-\infty, 0] & \text{when } a_i < f_i(x) = b_i, \\ [0, \infty) & \text{when } a_i = f_i(x) < b_i, \\ (-\infty, \infty) & \text{when } a_i < f_i(x) < b_i. \end{cases} \tag{2.11}$$

When I_i is a singleton $\{c_i\}$ designating an equality constraint $f_i(x) = c_i$, the interval in question is $T_{I_i}\big(f_i(x)\big) = \{0\}$.

Theorem 3. *In Example 2, consider any $x \in C$ at which the constraint qualification (1.5) is satisfied. Then*

$$\begin{aligned} \partial\delta_C(x) = N_C(x) &= \Big\{ \textstyle\sum_{i=1}^m y_i \nabla f_i(x) \,\Big|\, y_i \in N_{I_i}\big(f_i(x)\big) \Big\} + N_X(x), \\ (\delta_C)'_x(\xi) &= \delta_{T_C(x)}(\xi) \\ &= \begin{cases} 0 & \text{if } \xi \in T_X(x) \text{ and } \langle \nabla f_i(x), \xi \rangle \in T_{I_i}\big(f_i(x)\big) \text{ for all } i, \\ \infty & \text{otherwise.} \end{cases} \end{aligned} \tag{2.12}$$

For any $v \in N_C(x)$ the second-order epi-derivatives of δ_C at x exist for v and are given by

$$(\delta_C)''_{x,v}(\xi) = \max_{y \in Y(x,v)} \Big\{ \textstyle\sum_{i=1}^m y_i \langle \xi, \nabla^2 f_i(x) \xi \rangle \Big\} + \delta_{\Xi(x,v)}(\xi), \tag{2.13}$$

where $Y(x,v)$ is a bounded, polyhedral set and $\Xi(x,v)$ is a polyhedral cone, namely

$$\begin{aligned} Y(x,v) &= \Big\{ y \,\Big|\, y_i \in N_{I_i}\big(f_i(x)\big),\ v - \textstyle\sum_{i=1}^m y_i \nabla f_i(x) \in N_X(x) \Big\}, \\ \Xi(x,v) &= \Big\{ \xi \in T_C(x) \,\Big|\, \langle v, \xi \rangle = 0 \Big\} \\ &= \Big\{ \xi \in T_X(x) \,\Big|\, \langle \nabla f_i(x), \xi \rangle \in T_{I_i}\big(f_i(x)\big) \text{ for all } i,\ \langle v, \xi \rangle = 0 \Big\}. \end{aligned} \tag{2.14}$$

Furthermore, the mapping $\partial\delta_C = N_C$ is proto-differentiable at x for v with

$$(\partial\delta_C)'_{x,v}(\xi) = \begin{cases} \left\{\sum_{i=1}^m y_i \nabla^2 f_i(x)\xi \,\middle|\, y \in Y_{\max}(x,v,\xi)\right\} + K_{x,v}(\xi) & \text{if } \xi \in \Xi(x,v), \\ \emptyset & \text{if } \xi \notin \Xi(x,v), \end{cases} \tag{2.15}$$

where $Y_{\max}(x,v,\xi)$ is the closed face of $Y(x,v)$ consisting of the multiplier vectors y that achieve the maximum in (2.13), and $K_{x,v}(\xi)$ is the polyhedral cone defined by

$$\begin{aligned} K_{x,v}(\xi) = &\left\{ \sum_{i=1}^m y_i \nabla f_i(x) \,\middle|\, y_i \in N_{I_i}\big(f_i(x)\big),\ y_i = 0 \text{ if } \big\langle \nabla f_i(x), \xi\big\rangle \neq 0 \right\} \\ &+ \left\{ z \in N_X(x) \,\middle|\, \langle z, \xi\rangle = 0 \right\} + \left\{ sv \,\middle|\, s \in I\!R \right\}. \end{aligned} \tag{2.16}$$

Proof. This time we apply Theorem 1 to the mapping F and function g in (1.6). We have in the notation $w = (u_1, \ldots, u_m, x) \in D = I_1 \times \cdots \times I_m \times X$ that

$$\begin{aligned} &\partial g(w) = N_D(w) = N_{I_1}(u_1) \times \cdots \times N_{I_m}(u_m) \times N_X(x), \\ &g'_w(\omega) = \delta_{T_D(w)}(\omega) \text{ with } T_D(w) = T_{I_1}(u_1) \times \cdots \times T_{I_m}(u_m) \times T_X(x). \end{aligned} \tag{2.17}$$

On the other hand, the transpose Jacobian matrix $\nabla F(x)^*$ has the gradients $\nabla f_i(x)$ as its first m columns, followed by the n columns of the $n \times n$ identity matrix, so that in the notation $y = (y_1, \ldots, y_m, z) \in I\!R^m \times I\!R^n$ we have $\nabla F(x)^* y = \sum_{i=1}^m y_i \nabla f_i(x) + z$. With these choices the claimed formulas follow immediately from the ones in Theorem 1, except for some work in identifying the cone $N_{\Xi(x,v)}(\xi)$ in Theorem 1 with the cone $K_{x,v}(\xi)$ in (2.16), which we now undertake.

From (2.14) we know that $\Xi(x,v)$ is the intersection of a certain family of polyhedral cones: $T_X(x)$, the subspace $H = \{\, \xi \,|\, \langle v, \xi\rangle = 0 \,\}$, and

$$K_i = \left\{ \xi \,\middle|\, \big\langle \nabla f_i(x), \xi \big\rangle \in T_{I_i}\big(f_i(x)\big) \right\} \text{ for } i = 1, \ldots, m.$$

The normal cone to $\Xi(x,v)$ at ξ is therefore the sum of the normal cones to each of these sets at ξ (no closure operation being necessary because of the polyhedral property):

$$N_{\Xi(x,v)}(\xi) = \sum_{i=1}^m N_{K_i}(\xi) + N_{T_X(x)}(\xi) + N_H(\xi).$$

For any convex cone K, the normal cone $N_K(\xi)$ at a vector $\xi \in K$ consists of the vectors h in the polar cone K^* such that $\langle h, \xi\rangle = 0$. We calculate through this that

$$N_{K_i}(\xi) = \begin{cases} \{\, y_i \nabla f_i(x) \,|\, y_i \in N_{I_i}\big(f_i(x)\big) \,\} & \text{if } \big\langle \nabla f_i(x), \xi \big\rangle = 0, \\ \{0\} & \text{if } \big\langle \nabla f_i(x), \xi \big\rangle \neq 0, \end{cases}$$

whereas

$$N_{T_X(x)}(\xi) = \Big\{ z \in N_X(x) \,\Big|\, \langle z, \xi \rangle = 0 \Big\}, \qquad N_H(\xi) = \Big\{ sv \,\Big|\, s \in I\!R \Big\}.$$

Thus, $N_{\Xi(x,v)}(\xi)$ is the same as the cone $K_{x,v}(\xi)$ described in (2.16). □

Note that although δ_C and $(\delta_C)'_x$ are indicator functions (having no values other than 0 and ∞), $(\delta_C)''_{x,v}$ is generally *not* an indicator function. Second-order epi-derivatives of δ_C can differ from 0 because they reflect the curvature properties of C. As a matter of fact, it is only in the case where C is polyhedral and therefore totally lacking in curvature that $(\delta_C)''_{x,v}$ is again an indicator.

For a simple illustration, suppose C is defined by a single $\mathcal{C}^2$ inequality constraint, $C = \{ x \,|\, f_1(x) \le 0 \}$, and $\bar{x}$ is a point where this is active. The condition takes the form $f_i(\bar{x}) \in I_1$ for $I_1 = (-\infty, 0]$, and we have $N_{I_1}\big(f_1(\bar{x})\big) = [0, \infty)$. The constraint qualification requires $\nabla f_1(\bar{x}) \neq 0$. The set $\partial \delta_C(\bar{x}) = N_C(\bar{x})$ consists of all nonnegative multiples of $\nabla f_1(\bar{x})$. In treating a particular element $\bar{v} = \bar{y}_1 \nabla f_1(\bar{x})$, we have to distinguish the cases where $\bar{y}_1 > 0$ or $\bar{y}_1 = 0$ (and therefore $\bar{v} = 0$). With $\bar{y}_1 > 0$, we get

$$(\delta_C)''_{\bar{x},\bar{v}}(\xi) = \begin{cases} \bar{y}_1 \big\langle \xi, \nabla^2 f_1(\bar{x}) \xi \big\rangle & \text{if } \big\langle \nabla f_1(\bar{x}), \xi \big\rangle = 0, \\ \infty & \text{if } \big\langle \nabla f_1(\bar{x}), \xi \big\rangle \neq 0, \end{cases}$$

$$(\partial \delta_C)'_{\bar{x},\bar{v}}(\xi) = \begin{cases} \big\{ \bar{y}_1 \nabla^2 f_1(\bar{x}) \xi + s \nabla f_1(\bar{x}) \,\big|\, s \in I\!R \big\} & \text{if } \langle \xi, \nabla f_1(\bar{x}) \rangle = 0, \\ \emptyset & \text{if } \langle \xi, \nabla f_1(\bar{x}) \rangle \neq 0. \end{cases}$$

On the other hand, with $\bar{y} = 0$ we get

$$(\delta_C)''_{\bar{x},\bar{v}}(\xi) = \begin{cases} 0 & \text{if } \big\langle \nabla f_1(\bar{x}), \xi \big\rangle \le 0, \\ \infty & \text{if } \big\langle \nabla f_1(\bar{x}), \xi \big\rangle > 0, \end{cases}$$

$$(\partial \delta_C)'_{\bar{x},\bar{v}}(\xi) = \begin{cases} \{0\} & \text{if } \big\langle \xi, \nabla f_1(\bar{x}) \big\rangle < 0, \\ \big\{ s \nabla f_1(\bar{x}) \,\big|\, s \ge 0 \big\} & \text{if } \big\langle \xi, \nabla f_1(\bar{x}) \big\rangle = 0, \\ \emptyset & \text{if } \big\langle \xi, \nabla f_1(\bar{x}) \big\rangle > 0. \end{cases}$$

Next we tackle Example 3. Our result in this situation is closely related to the one for Example 2, but to bring out connections with nonlinear programming theory we state it in terms of the Lagrangian function

$$L(x, y) = f_0(x) + y_1 f_1(x) + \cdots + y_m f_m(x).$$

Theorem 4. *In Example 3, consider any $x \in C$ at which the constraint qualification (1.5) is satisfied. Then*

$$\begin{aligned} \partial f(x) &= \nabla f_0(x) + N_C(x) = \Big\{ \nabla_x L(x, y) \,\Big|\, y_i \in N_{I_i}\big(f_i(x)\big) \Big\} + N_X(x), \\ f'_x(\xi) &= \begin{cases} \big\langle \nabla f_0(x), \xi \big\rangle & \textit{if } \xi \in T_X(x) \textit{ and } \big\langle \nabla f_i(x), \xi \big\rangle \in T_{I_i}\big(f_i(x)\big) \textit{ for all } i, \\ \infty & \textit{otherwise.} \end{cases} \end{aligned} \tag{2.18}$$

For any $v \in \partial f(x)$ the second-order epi-derivatives of f at x exist for v and are given in terms of the Lagrangian L by

$$f''_{x,v}(\xi) = \max_{y \in Y(x,v)} \Big\langle \xi, \nabla^2_{xx} L(x,y)\xi \Big\rangle + \delta_{\Xi(x,v)}(\xi), \tag{2.19}$$

where $Y(x,v)$ is a bounded, polyhedral set and $\Xi(x,v)$ is a polyhedral cone, namely

$$\begin{aligned} Y(x,v) &= \Big\{ y \,\Big|\, y_i \in N_{I_i}\big(f_i(x)\big),\ v - \nabla_x L(x,y) \in N_X(x) \Big\}, \\ \Xi(x,v) &= \Big\{ \xi \in T_C(x) \,\Big|\, \langle v - \nabla f_0(x), \xi \rangle = 0 \Big\} \\ &= \Big\{ \xi \in T_X(x) \,\Big|\, \langle \nabla f_i(x), \xi \rangle \in T_{I_i}\big(f_i(x)\big) \text{ for all } i,\ \langle v - \nabla f_0(x), \xi \rangle = 0 \Big\}. \end{aligned} \tag{2.20}$$

Furthermore, the mapping ∂f is proto-differentiable at x for v with

$$(\partial f)'_{x,v}(\xi) = \begin{cases} \Big\{ \nabla^2_{xx} L(x,y)\xi \,\Big|\, y \in Y_{\max}(x,v,\xi) \Big\} + K_{x,v}(\xi) & \text{if } \xi \in \Xi(x,v), \\ \emptyset & \text{if } \xi \notin \Xi(x,v), \end{cases} \tag{2.21}$$

where $Y_{\max}(x,v,\xi)$ is the closed face of $Y(x,v)$ consisting of the multiplier vectors y that achieve the maximum in (2.19), and $K_{x,v}(\xi)$ is the polyhedral cone defined by

$$\begin{aligned} K_{x,v}(\xi) =& \Big\{ \textstyle\sum_{i=1}^m y_i \nabla f_i(x) \,\Big|\, y_i \in N_{I_i}\big(f_i(x)\big) \text{ but } y_i = 0 \text{ if } \langle \nabla f_i(x), \xi \rangle \neq 0 \Big\} \\ & + \Big\{ z \in N_X(x) \,\big|\, \langle z, \xi \rangle = 0 \Big\} + \Big\{ s\big[v - \nabla f_0(x)\big] \,\Big|\, s \in \mathbb{R} \Big\}. \end{aligned} \tag{2.22}$$

Proof. These facts can be derived in close parallel with the ones in Theorem 3, to which they specialize when $f_0 \equiv 0$. The composite representation in (1.8) serves this purpose, but an alternative approach is to observe that because f_0 is a $\mathcal{C}^2$ function the second epi-derivatives of f in this case can be obtained from the ones in Theorem 3 merely by the addition of the term $\langle \xi, \nabla^2 f_0(x)\xi \rangle$, and with v replaced by $v - \nabla f_0(x)$ in the formulas for $Y(x,v)$ and $\Xi(x,v)$. Then the proto-derivatives of $\partial f(x)$ can be obtained similarly by adding the term $\nabla^2 f_0(x)\xi$ to the formula in Theorem 3 and replacing v by $v - \nabla f_0(x)$ in the formula for $K_{x,v}(\xi)$. $\square$

In Example 3 and Theorem 4 the function f_0 has been assumed to be $\mathcal{C}^2$, but the methodology is not limited to that case. We could easily go further by taking $f = f_0 + \delta_C$ with the set C chosen according to the specifications in Example 2, but with f_0 taken to be any fully amenable function. In particular, f_0 could be a max function of the kind in Example 1, hence nonsmooth. This generality is attained through the calculus we have developed in [8], which provides formulas for $f''_{x,v}(\xi)$ and $(\partial f)'_{x,v}(\xi)$ when f is expressed as the sum of two fully amenable functions under an associated "constraint qualification" on the domains of the functions. For $f = f_0 + \delta_C$ this constraint qualification is satisfied in particular when f_0 is finite everywhere, as in the max function case. Then $\partial f(x) = \partial f_0(x) + N_C(x)$, and for any $v \in \partial f(x)$ one has in terms of the set

$$V(x,v) := \big\{(v_0,v_1) \,\big|\, v_0 \in \partial f_0(x),\ v_1 \in N_C(x),\ v_0+v_1 = v\big\}$$

the expressions

$$\begin{aligned} f''_{x,v}(\xi) &= \max_{(v_0,v_1)\in V(x,v)} \Big\{(f_0)''_{x,v_0}(\xi) + (\delta_C)''_{x,v_1}(\xi)\Big\}, \\ (\partial f)'_{x,v}(\xi) &= \bigcup_{(v_0,v_1)\in V_{\max}(x,v,\xi)} \Big\{\big(\partial f_0\big)'_{x,v_0}(\xi) + \big(\partial \delta_C\big)'_{x,v_1}(\xi)\Big\}, \end{aligned} \tag{2.23}$$

where $V_{\max}(x,v,\xi)$ is the set of vectors (v_0,v_1) that achieve the maximum in (2.23).

3. Comparison with Other Work

Recall that ∂f is proto-differentiable at x for $v \in \partial f(x)$ if

$$\limsup_{t\searrow 0} \big[\operatorname{gph}\partial f - (x,v)\big]/t = \liminf_{t\searrow 0} \big[\operatorname{gph}\partial f - (x,v)\big]/t$$

(where gph stands for graph), and note that

$$(\xi,z) \in \limsup_{t\searrow 0} \big[\operatorname{gph}\partial f - (x,v)\big]/t \quad\Longleftrightarrow\quad z \in \limsup_{\substack{\xi'\to\xi \\ t\searrow 0}} \big[\partial f(x+t\xi') - v\big]/t.$$

As mentioned in the Introduction, proto-derivatives of ∂f for the max functions f in Example 1 have been studied by Auslender and Cominetti [12] and Penot [13]. In [12] and [13] the following formula is given for the outer graphical limit of the difference quotients: For any $v \in \partial f(x)$,

$$\limsup_{\substack{\xi'\to\xi \\ t\searrow 0}} \big[\partial f(x+t\xi') - v\big]/t = \bigcup_{I^*\in S(x,v,\xi)} \ \bigcup_{y\in Y(I^*,v)} \Big[\textstyle\sum_{i=1}^m y_i \nabla^2 f_i(x)\xi + E(I^*,y)\Big], \tag{3.1}$$

where

$$\begin{aligned} Y(I^*,v) &:= \big\{\, y \in Y(x,v) \,\big|\, y_i = 0 \ \text{ for } \ i \notin I^* \,\big\}, \\ S(x,v,\xi) &:= \big\{I^* \subset I(x) \,\big|\, Y(I^*,v) \neq \emptyset \ \text{ and } \ \exists\, t_k \searrow 0,\ \xi_k \to \xi, \\ &\qquad\qquad\qquad\qquad \text{with } \ I^* = I(x+t_k\xi_k) \ \text{ for all } \ k\big\}, \\ E(I^*,y) &:= \big\{\ \textstyle\sum_{i=1}^m \sigma_i \nabla f_i(x) \,\big|\, \sum_{i=1}^m \sigma_i = 0,\ \sigma_i = 0 \ \text{ if } \ i \notin I^*,\ \sigma_i \geq 0 \ \text{ if } \ y_i = 0 \,\big\}. \end{aligned}$$

It was observed by Cominetti and Auslender that an element I^* of $S(x,v,\xi)$ is actually included in $I(x,\xi)$ (the set of indices achieving the maximum in (2.6)). With this observation one can easily check that if $S(x,v,\xi)$ is nonempty, then ξ must be an element of $\Xi(x,v)$. Furthermore in [13], under the condition that

$$\liminf_{t \searrow 0} \big[M(I(x,\xi)) - x\big]/t = \big\{\, \xi' \in I\!R^n \,\big|\, \langle \nabla f_i(x), \xi' \rangle = f'_x(\xi') \quad \forall i \in I(x,\xi) \,\big\}, \tag{3.2}$$

where $M\big(I(x,\xi)\big) := \big\{\, x' \in I\!R^n \,\big|\, f_i(x') = f(x'), \forall i \in I(x,\xi)\big\}$, then ∂f is proto-differentiable at x for v "in the direction ξ" (see [13]) with

$$(\partial f)'_{x,v}(\xi) = \bigcup_{y \in Y(I(x,\xi),v)} \big[\textstyle\sum_{i=1}^m y_i \nabla^2 f_i(x)\xi + E(I(x,\xi),y)\big]. \tag{3.3}$$

Finally, in both [12] and [13] the special case of $\big\{\, \nabla f_i(x) \,\big|\, i \in I(x,\xi) \,\big\}$ linearly independent is discussed (the case $\big\{\, \nabla f_i(x) \,\big|\, i \in I(x,\xi) \,\big\}$ affinely independent is the only example satisfying condition (3.2) that is provided in [13]). The formula for the proto-derivatives in this case becomes

$$(\partial f)'_{x,v}(\xi) = \textstyle\sum_{i=1}^m y_i \nabla^2 f_i(x)\xi + E\big(I(x,\xi),y\big), \tag{3.4}$$

where $\xi \in \Xi(x,v)$ while y is the unique element of $Y(x,v)$.

In order to simplify formulas (3.1) and (3.3), and to compare the various formulas with each other, we need to give an alternate description of the cones $K_{x,v}(\xi)$ and $E(I^*,y)$.

Proposition 1. *Fix $\xi \in \Xi(x,v)$. For any $y \in Y(x,v)$,*

$$K_{x,v}(\xi) = E\big(I(x,\xi),y\big). \tag{3.5}$$

Furthermore, for any index set $I^* \in S(x,v,\xi)$ and any vector $y \in Y(I^*,v)$, the polyhedral cone $E(I^*,y)$ is the convex cone generated by the vectors $\nabla f_i(x) - v$ for $i \in I^*$. In particular, $E(I^*,y) \subset K_{x,v}(\xi)$.

Proof. Fix $\xi \in \Xi(x,v)$, and $y \in Y(x,v)$. Because

$$\big\langle v,\xi\big\rangle = \sum_{i=1}^m y_i \big\langle \nabla f_i(x),\xi\big\rangle = \max_{i \in I(x)} \big\langle \nabla f_i(x),\xi\big\rangle,$$

it follows that $y_i = 0$ if $i \notin I(x,\xi)$. So actually $v = \sum_{i \in I(x,\xi)} y_i \nabla f_i(x)$. If $w \in K_{x,v}(\xi)$ then, by Theorem 2, w is in the convex cone generated by the vectors $\nabla f_i(x) - v$ for

$i \in I(x,\xi)$. We then have

$$\begin{aligned} w &= \sum_{i \in I(x,\xi)} \mu_i \Big(\nabla f_i(x) - v\Big) \text{ with } \mu_i \geq 0 \\ &= \sum_{i \in I(x,\xi)} \mu_i \left(\nabla f_i(x) - \sum_{j \in I(x,\xi)} y_j \nabla f_j(x)\right) \text{ with } \mu_i \geq 0 \\ &= \sum_{i \in I(x,\xi)} \left(\mu_i - y_i \sum_{j \in I(x,\xi)} \mu_j\right) \nabla f_i(x) \text{ with } \mu_i \geq 0. \end{aligned}$$

Let $\sigma_i = \big(\mu_i - y_i \sum_{j \in I(x,\xi)} \mu_j\big)$. Then $\sigma_i \geq 0$ if $y_i = 0$, and $\sum_{i \in I(x,\xi)} \sigma_i = 0$, i.e., w is in $E\big(I(x,\xi), y\big)$. To establish the reverse inclusion in (3.5), take w in $E\big(I(x,\xi), y\big)$ and let $\mu_i = \sigma_i + y_i M$, where $M := \max\{-\sigma_i / y_i \mid y_i \neq 0\}$. It follows that $\mu_i \geq 0$ and $\sum_{i \in I(x,\xi)} \mu_i (\nabla f_i(x) - v) = w$, i.e., w is in $K_{x,v}(\xi)$. By a similar argument $E(I^*, y)$ is the convex cone generated by the vectors $\nabla f_i(x) - v$ for $i \in I^*$. ☐

In the special case with $\{\, \nabla f_i(x) \,|\, i \in I(x,\xi) \,\}$ linearly independent, it is easy to see that (3.4) agrees with (2.9). In other cases the reconciliation of (3.1) with (2.9) is not a simple task. One difficulty in comparing the formulas is that it's hard in the much more complicated framework of (3.1) to identify just which index sets I^* belong to the collection $S(x, v, \xi)$, a circumstance that led Cominetti and Auslender to comment that the computation of $S(x, v, \xi)$ can only be carried out in special situations. There is no such obstacle in applying our formula (2.9).

Another difficulty, illustrated by the following example, is that $E(I^*, y)$ can sometimes be a proper subset of $K_{x,v}(\xi)$. Consider the function $f(x_1, x_2) := \max_{i \in \{1,2,3\}} f_i(x)$, where $f_1(x) = \frac{1}{2} x_1^2$, $f_2(x) = x_2$, and $f_3(x) = -x_2$, at the points $x = (0,0)$, $v = (0,0)$, and $\xi = (1,0)$. A simple calculation shows that $Y(x,v) = \{\, (1-2\alpha, \alpha, \alpha) \,|\, 0 \leq \alpha \leq 1/2 \,\}$, $K_{x,v}(\xi) = \{\, (\xi_1, \xi_2) \,|\, \xi_1 = 0 \,\}$, and $S(x, v, \xi) = \big\{\{1,2\}, \{1,3\}, \{1\}\big\}$, with $E\big(\{1,2\}, (1,0,0)\big) = \{\, \lambda(0,1) \,|\, \lambda \geq 0 \,\}$, and $E\big(\{1,3\}, (1,0,0)\big) = \{\, \lambda(0,1) \,|\, \lambda \leq 0 \,\}$. Nevertheless, the two formulas (2.9) and (3.1) are confirmed as agreeing in this example.

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. A. Poliquin and R. T. Rockafellar, "Amenable functions in optimization," to appear in the Proceedings of the International School of Mathematics G. Stampacchia, 10th course: Nonsmooth Optimization: Methods and Applications.
3. R. T. Rockafellar, "Nonsmooth analysis and parametric optimization," in *Methods of Nonconvex Analysis* (A. Cellina, ed.), Springer-Verlag Lecture Notes in Math. No. 1446 (1990), 137–151.
4. R. T. Rockafellar, "Perturbation of generalized Kuhn-Tucker points in finite dimensional optimization," in *Nonsmooth Optimization and Related Topics*, (F.H. Clarke et al. eds.), Plenum Press, 1989, 393–402.
5. R. A. Poliquin, "Proto-differentiation of subgradient set-valued mappings," Canadian J. Math. **42** (1990), 520–532.
6. R. T. Rockafellar, "First- and second-order epi-differentiability in nonlinear programming," Trans. Amer. Math. Soc. **307** (1988), 75–107.
7. R. T. Rockafellar, "Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives," Math. Oper. Research **14** (1989), 462–484.
8. R. A. Poliquin and R. T. Rockafellar, "A calculus of epi-derivatives applicable to optimization," Canadian J. Math., to appear.
9. F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, 1983.
10. B. S. Mordukhovich, *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow (1988) (in Russian); English translation to appear in Wiley-Interscience.
11. R. T. Rockafellar, "Lagrange multipliers and optimality," SIAM Review, to appear.
12. A. Auslender and R. Cominetti, "A comparative study of multifunction differentiability with applications in mathematical programming," Math. of Oper. Research **16** (1991), 240–258.
13. J.P. Penot, "On the differentiability of the subdifferential of a maximum function," preprint, 1992.
14. R. T. Rockafellar, "Maximal monotone relations and the second derivatives of nonsmooth functions," Ann. Inst. H. Poincaré: Analyse non linéaire **2** (1985), 167–184.
15. R. T. Rockafellar, "Generalized second derivatives of convex functions and saddle functions," Trans. Amer. Math. Soc. **322** (1990), 51–77.

16. C. Do, "Generalized second derivatives of convex functions in reflexive Banach spaces," Trans. Amer. Math. Soc. **334** (1992), 281–301.