AMENABLE FUNCTIONS IN OPTIMIZATION

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

1. INTRODUCTION

Any optimization problem on $I\!R^n$ can be formulated as the minimization of an extended real-valued function f over all of $I\!R^n$ (i.e., $f : I\!R^n \to \overline{I\!R}$, where $\overline{I\!R} := I\!R \cup \{\infty\}$). In this setting f is called the *essential* objective function for the problem, and the effective domain $\operatorname{dom} f := \{x \mid f(x) < \infty\}$ represents the set of feasible solutions.

The idea of considering optimization problems where the function f belongs to a specific class of functions e.g. locally Lipschitzian, convex, smooth, is of course not new to optimization theory. Classical optimization dealt with smooth functions f (with no constraints), while the theory of convex optimization developed in the 60's concentrated on the case of f convex. Eventually the idea dawned that many, if not most, of the problems commonly seen in applications and solved on computers corresponded to convexly composite functions f, where $f(x) = g\big(F(x)\big)$ for an extended-real-valued convex function g and a smooth mapping F, cf. [1]–[7]. This includes standard nonlinear programming problems (even with constraints), convex problems, nonsmooth (and nonconvex) problems in which the objective is the max of a finite collection of smooth functions, and much more. Further, functions f of such type enjoy special properties of strong interest in the statement of optimality conditions and the design of algorithms, but which are not well captured by classical analysis, convex analysis, or general forms of nonsmooth analysis.

In light of such facts it appears important to study well chosen classes of convexly

---

* Dept. of Math., Univ. of Alberta, Edmonton, Canada. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under grant OGP41983.

** Dept. of Math., Univ. of Washington, Seattle, USA. This work was supported in part by the National Science Foundation under grant DMS–8819586.

composite functions, viewing their properties independently of any particular representation $f = g \circ F$, which after all need not be unique. Some work in this direction has already been undertaken in [7], where a calculus of first and second derivatives has been developed which is much sharper than anything available for general convex or nonsmooth functions. The purpose of the present paper is to provide some perspective to those results by showing how they can be applied to a broad problem model in nonconvex optimization. Results on second-order optimality conditions and on perturbations of optimal solutions will be obtained, among others.

The following definitions help us free ourselves of the burden of always having to refer to a specific representation in terms of a convex function and a smooth mapping, when this is immaterial and might even get in the way of appreciating the generality of the class of functions that is involved.

Definition 1.1. *A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ 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 $\mathcal{C}^1$ mapping $F : V \to \mathbb{R}^d$ and a proper, lower semicontinuous, convex function $g : \mathbb{R}^d \to \overline{\mathbb{R}}$ such that $f(x) = g\big(F(x)\big)$ for $x \in V$ and*

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

Here $\nabla F(\bar{x})$ denotes the $d \times n$ Jacobian matrix of F at $\bar{x}$, and $\nabla F(\bar{x})^*$ is its transpose. Further, $N\big(F(\bar{x}) \big| \operatorname{dom} g\big)$ is the normal cone to the nonempty convex set $\operatorname{dom} g$ at the point $F(\bar{x})$ in the sense of convex analysis. Condition (1.1) has the role of a basic constraint qualification in situations where f is the essential objective in an optimization problem, since the condition $F(x) \in \operatorname{dom} g$, which defines $\operatorname{dom} f$, then determines the feasible solutions.

Definition 1.2. *A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is fully amenable at $\bar{x}$ if the conditions in Definition 1.1 can be satisfied with the extra stipulation that F is a $\mathcal{C}^2$ mapping and g is piecewise linear-quadratic (convex). The latter means that $\operatorname{dom} g$ can be expressed as the union of a finite collection of polyhedral (convex) sets, on each of which g is given by a polynomial function with no terms higher than degree two.*

Any lower semicontinuous, proper, convex function is obviously amenable everywhere (meaning, at every point of its effective domain): just take F to be the identity. In particular the indicator of a closed, convex set is amenable. If the Jacobian mapping $\nabla F(\bar{x})$ is nonsingular, then for any choice of a lower semicontinuous, proper, convex function g, the function $f = g \circ F$ is amenable at $\bar{x}$; this is the case of a function f that is "convex except for a poor choice of coordinates." Other examples of amenable functions, not fitting this simple view of hidden convexity, are $\mathcal{C}^1$ functions and all functions expressible as the pointwise maximum of finitely many $\mathcal{C}^1$ functions; see [7].

Full amenability is the refinement of amenability that supports second-order as well as first-order theory, as will be seen below. Polyhedral functions (having polyhedral epigraph) are examples of piecewise linear-quadratic (convex) functions. Specific instances would be the indicator function δ_C and support function σ_C of a polyhedral set C. A convex function is piecewise linear-quadratic if and only if its subdifferential mapping is polyhedral in the sense of Robinson [8], i.e., has a union of finitely many polyhedral sets as its graph, cf. Sun [9]. The conjugate of a convex, piecewise linear-quadratic function is again piecewise linear-quadratic [9].

A function of class $\mathcal{C}^2$ is obviously fully amenable everywhere, and so too is the pointwise maximum of finitely many $\mathcal{C}^2$ functions, because such a function can be viewed as the composition of a $\mathcal{C}^2$ mapping with a simple piecewise linear convex function (the support function of the unit simplex). In the same vein, a set given by finitely many C^2 constraints is seen to be fully amenable at points where the Mangasarian-Fromovitz constraint qualification is satisfied, and likewise a function obtained by adding the indicator of such a set to a max of finitely many $\mathcal{C}^2$ functions is fully amenable at such points, cf. [7]. In particular, a function that the sum of a C^2 function and the indicator of a polyhedral set is fully amenable at all points of its effective domain. Such examples emerge from the choice of particular composite representations, but still others follow from application of the calculus rules in [7] (and below).

The problem model we turn to in this paper is

$$(\mathcal{P}) \qquad \text{minimize} \quad k(x) + h\big(G(x)\big) \ \text{ over all } \ x \in \mathbb{R}^n,$$

with $k : \mathbb{R}^n \to \overline{\mathbb{R}}$, $G : \mathbb{R}^n \to \mathbb{R}^m$, and $h : \mathbb{R}^m \to \overline{\mathbb{R}}$. By the *amenable case* of $(\mathcal{P})$, we shall mean the case where k and h are amenable everywhere (on their effective domains) and G is $\mathcal{C}^1$. The *fully amenable case*, for the sake of obtaining second-order results, will be the one where k and h are fully amenable everywhere and G is $\mathcal{C}^2$. The essential objective function in $(\mathcal{P})$ is of course $f = k + h \circ G$. The feasible solutions are the points x such that $f(x) < \infty$, or in other words, such that

$$x \in \operatorname{dom} k, \qquad G(x) \in \operatorname{dom} h. \tag{1.2}$$

An important special case of problem $(\mathcal{P})$, serving well to establish the centrality of the model in optimization, is obtained when h is the indicator function δ_D of some set $D \subset \mathbb{R}^m$. The set D is said to be amenable if δ_D is amenable (e.g. if D is closed and convex) and fully amenable if δ_D is fully amenable (e.g. if D is polyhedral). Then $(\mathcal{P})$ has the form

$$\text{minimize } k(x) \text{ subject to } G(x) \in D. \tag{1.3}$$

In the general case of $(\mathcal{P})$ one can think of the term $h\big(G(x)\big)$ as possibly being a penalty expression for the relation between $G(x)$ and some target set, the penalties being finite instead of infinite, cf. [10].

It will be essential to have an appropriate constraint qualification for working with problem $(\mathcal{P})$ and the system (1.2), and the condition in Definition 1.1 can serve as a guide. A crucial difference, however, between (1.2) and the situation in Definition 1.1 is that *the sets* $\operatorname{dom} k$ *and* $\operatorname{dom} h$ *might not be convex, and the function* h *might not be convex.* We therefore rely on the following extensions of the concepts of normal cone and subgradient, cf. Mordukhovich [11]. These are known to reduce to the usual concepts of convex analysis when the set or function happens to be convex, cf. Clarke [12] [13].

The *normal cone* to a closed (but not necessarily convex) set C at $\bar{x}$, denoted by $N\big(\bar{x}|C\big)$, is the cone (not necessarily convex) consisting of the zero vector and all the vectors v for which there exists a sequence of points $x^\nu \notin C$ (with $\nu = 1, 2, \ldots$) having nearest point projections $\bar{x}^\nu \in C$ with $\bar{x}^\nu \to \bar{x}$, such that $\lambda^\nu(x^\nu - \bar{x}^\nu) \to v$ for some choice of scalars $\lambda^\nu > 0$. For a function $f : I\!R^n \to \overline{I\!R}$ a vector $v \in I\!R^n$ is a *subgradient* of f at $\bar{x}$, denoted by $v \in \partial f(\bar{x})$, if $f(\bar{x})$ is finite and $(v, -1)$ belongs to the normal cone $N\big(\bar{x}, f(\bar{x}) \,\big|\, \operatorname{epi} f\big)$.

Definition 1.3. *A feasible solution $\bar{x}$ to $(\mathcal{P})$ will be said to satisfy the basic constraint qualification if*

$$\textit{there is no } y \neq 0 \textit{ in } N\big(G(\bar{x})|\operatorname{dom} h\big) \textit{ with } -\nabla G(\bar{x})^* y \in N(\bar{x}|\operatorname{dom} k). \tag{1.4}$$

Example 1.4. As an illustration of the basic constraint qualification, consider the mathematical programming problem

$$(\mathcal{P}_0) \qquad \text{minimize } f_0(x) \text{ over all } x \in C \text{ satisfying } f_i(x) \in I_i \text{ for } i = 1, \ldots, m,$$

where the functions $f_i : I\!R^n \to I\!R$ for $i = 0, 1, \ldots, m$, are of class $\mathcal{C}^2$, the nonempty set C is polyhedral (implying convex), and each nonempty set I_i is a closed (but not necessarily bounded) interval in $I\!R$. This is a fully amenable case of $(\mathcal{P})$ corresponding to

$$\begin{aligned} &k(x) = f_0(x) + \delta_C(x), \qquad G(x) = \big(f_1(x), \ldots, f_m(x)\big) \in I\!R^m, \\ &h(x) = \delta_D(x) \text{ with } D = I_1 \times \ldots \times I_m \subset I\!R^m. \end{aligned} \tag{1.5}$$

In this case the basic constraint qualification (1.4) can be worked out without appeal to the generalized definitions of normal cone and subgradient. It takes the form:

$$\begin{cases} \text{there is no } (y_1,\ldots,y_m) \neq (0,\ldots,0) \text{ with} \\ y_i \in N\big(f_i(\bar{x})\big|I_i\big) \text{ and } -\sum_{i=1}^m y_i \nabla f_i(\bar{x}) \in N\big(\bar{x}|C\big). \end{cases} \tag{1.6}$$

Here if $I_i = [0, -\infty)$, for instance, then

$$N\big(f_i(\bar{x})\big|I_i\big) = \begin{cases} (-\infty, 0] & \text{if } f_i(\bar{x}) = 0 \\ 0 & \text{if } f_i(\bar{x}) > 0, \end{cases} \tag{1.7}$$

and similarly for other cases, such as $I_i = [a_i, b_i]$ for finite lower and upper bounds a_i and b_i. The basic constraint qualification (1.4) can be seen therefore as a generalization of the familiar Mangasarian-Fromovitz constraint qualification (dual version).

The plan for the remainder of the paper is as follows. In Section 2 we briefly review the notions of first- and second-order epi-differentiation and proto-differentiation. We also review there the applications of epi-derivatives to optimality conditions, underlining the need for identifying classes of twice epi-differentiable functions. In Section 3 we recall the principal calculus results of [7] and employ them in deriving second-order optimality conditions for problem $(\mathcal{P})$. In Section 4 we present an application of the calculus results to perturbation of solutions in parametric optimization.

2. EPI- AND PROTO-DIFFERENTIATION

In order to lay the groundwork for stating the calculus results in Section 3, we need to review the notions of first- and second-order epi-differentiation of a function and proto-differentiation of a set-valued mapping. We begin with epi-derivatives.

The study of limits of difference quotients is at the heart of any kind of differentiability theory, but usually the limits are taken in a pointwise sense. Epi-derivatives depend instead on difference quotient functions converging instead in the epigraphical sense. A sequence $\{g_t\}_{t>0}$ of extended-real-valued functions on $I\!R^n$ is said to *epi-converge* to g if the epigraphs $\operatorname{epi} g_t = \{(x,\alpha) \in I\!R^n \times I\!R : \; g_t(x) \geq \alpha\}$ converge as sets; see [14], [15] and [16] for a discussion of epi-convergence and its significance in optimization theory.

Recall that for a sequence $\{C_t\}_{t>0}$, of nonempty sets in $I\!R^n$, the set $\liminf_{t \searrow 0} C_t$ consists of all points $\lim_{t\searrow 0} c_t$ obtainable from a selection of $c_t \in C_t$ (for all $t > 0$ sufficiently small), whereas the set $\limsup_{t\searrow 0} C_t$ consists of all cluster points obtainable from such a selection (to the extent that $C_t \neq \emptyset$). One says that C_t converges in the Painlevé-Kuratowski sense to C if

$$\liminf_{t\searrow 0} C_t = \limsup_{t\searrow 0} C_t = C.$$

First- and second-order epi-derivatives were introduced in Rockafellar [5] and developed further in Rockafellar [6],[17],[18], Cominetti [19], Do [20], Poliquin [21],[22], and Poliquin and Rockafellar [23]. A lower semi-continuous function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be *epi-differentiable* at a point x where $f(x)$ is finite if the first-order difference quotient functions $\Delta_{x,t} f : \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by

$$\Delta_{x,t} f(x') = \big[f(x + tx') - f(x)\big]/t \ \text{ for } \ t > 0$$

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. We say that v is an *epi-gradient* of f at x if $f'_x(x') \geq \langle v, x' \rangle$ for all x'. Epi-gradients are quite familiar objects in nonsmooth analysis; if v is an epi-gradient of f at x then

$$f(x + x') \geq f(x) + \langle v, x' \rangle + o(x'),$$

i.e., every epi-gradient is a lower semigradient (also called a Dini subgradient; see [24]). The converse is not true: a function f with a lower semigradient at x need not be epi-differentiable at x.

In terms of epigraphs, epi-differentiability means that

$$\operatorname{epi} f'_x \;=\; \lim_{t \searrow 0} \big\{ \big[\operatorname{epi} f - (x, f(x)) \big]/t \big\}. \tag{2.1}$$

Thus, f is epi-differentiable at x if and only if the *(Bouligand) contingent cone* (the cone defined by taking limsup in (2.1)) exists *as a limit.* In particular, if f is *Clarke regular* at x in the sense that the contingent cone to $\operatorname{epi} f$ at $(x, f(x))$ equals the Clarke tangent cone there, then f is epi-differentiable.

A function f is *twice epi-differentiable* at x relative to a vector $v \in \mathbb{R}^n$ if it is epi-differentiable at x and the second-order difference quotient functions $\Delta^2_{x,v,t} f : \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by

$$\Delta^2_{x,v,t} f(x') = \big[f(x + tx') - f(x) - t\langle v, x' \rangle\big]/\tfrac{1}{2}t^2 \ \text{ for } \ t > 0$$

epi-converge to a proper function as $t \searrow 0$. (It follows quite easily then that the vector v must be an epi-gradient of f at x.) The limit function is the second-order epi-derivative, denoted by $f''_{x,v}(x')$.

A simple example occurs when the function f happens to be of class $\mathcal{C}^2$. Then the second-order epi-derivative exists and is given by:

$$f''_{x,v}(x') \;=\; \langle x', \nabla^2 f(x) x' \rangle \ \text{ for } \ v = \nabla f(x),$$

where $\nabla^2 f(x)$ is the Hessian of f at x.

Optimality conditions resembling the classical ones for minimizing smooth functions have been derived in Rockafellar [6, Thm. 2.2] for a twice epi-differentiable function f.

Necessary conditions: If $\bar{x}$ furnishes a local minimum of f, then $f'_x(x') \geq 0$ for all x' (equivalently, $0 \in \operatorname{coder} f(\bar{x})$) and $f''_{x,0}(x') \geq 0$ for all x'.

Sufficient conditions: If $\bar{x}$ is a point where $f'_x(x') \geq 0$ for all x' (equivalently, $0 \in \operatorname{coder} f(\bar{x})$) and $f''_{x,0}(x') > 0$ for all $x' \neq 0$, then $\bar{x}$ furnishes a local minimum of f in the strong sense.

(The "strong sense" means that $f(x) \geq f(\bar{x}) + \varepsilon|x - \bar{x}|^2$ locally for some $\varepsilon > 0$.) Obviously, the issue remains of determining what the derivatives in question actually are in a given case, but this is precisely why it is important to have a calculus available, such as will be described in the next section. The existence of such a calculus is a favorable feature of epi-derivatives of amenable functions. Optimality conditions can be developed under less demanding assumptions, cf. Ioffe [4], but they may require more effort in their application and yield conclusions that are not as potent. The philosophy behind epi-differentiability is to trace this sharper case as far as possible before widening the scope of the theory, since many of the problems that come up do naturally fall into this case anyway.

Next we recall the notion of proto-differentiation. A set-valued mapping $\Gamma : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *proto-differentiable* at x relative to the element $v \in \Gamma(x)$ if the (set-valued) difference quotient mappings

$$\big(\Delta_{x,v,t}\Gamma\big)(x') = \big[\Gamma(x + tx') - v\big]/t$$

graph-converge as $t \searrow 0$. (This means that their graphs, as subsets of $\mathbb{R}^n \times \mathbb{R}^n$ indexed by t, have a limit as $t \searrow 0$.) If so, the limit mapping is denoted by $\Gamma'_{x,v}$ and called the *proto-derivative* of Γ at x relative to v. This proto-derivative mapping assigns to each $x' \in \mathbb{R}^n$ a (possibly empty) subset $\Gamma'_{x,v}(x')$ of $\mathbb{R}^n$.

Many of the set-valued mappings encountered in optimization, e.g. set-valued mappings expressing feasibility or optimality, are proto-differentiable; see [25].

When applied to a subgradient mapping $\operatorname{coder} g$, proto-differentiation gives rise to a second-order theory for g. For a convex function g we only need to invoke Attouch's Theorem (see [17]) to show that the proto-derivative of the subgradient mapping $\operatorname{coder} g$ is related to the second-order epi-derivative $g''_{x,v}$ by the equation:

$$(\operatorname{coder} g)'_{x,v}(x') = \operatorname{coder}\big(\tfrac{1}{2} g''_{x,v}\big)(x') \quad \text{for all } x'.$$

As we shall see in the next section, this formula also holds for fully amenable functions, despite their lack of convexity in general. In the setting of parametric optimization the formula will be used to show that the proto-derivatives of the solution mapping can be obtained as primal-dual pairs for an auxiliary derivative problem; see Section 4.

3. CALCULUS AND OPTIMALITY CONDITIONS

The initial theorem in this section summarizes some of the more important first- and second-order properties of amenable functions.

Theorem 3.1 (cf. [7, Theorem 2.9]). *If f is amenable at $\bar{x}$, then f is both epi-differentiable and Clarke regular at $\bar{x}$ with*

$$\begin{aligned} \partial f(\bar{x}) &= \Big\{ v \,\Big|\, \langle v, x' \rangle \leq f'_{\bar{x}}(x') \text{ for all } x' \Big\}, \\ f'_{\bar{x}}(x') &= \sup \Big\{ \langle v, x' \rangle \,\Big|\, v \in \operatorname{coder} f(\bar{x}) \Big\}. \end{aligned} \tag{3.1}$$

If f is fully amenable at $\bar{x}$, it is twice epi-differentiable at $\bar{x}$ relative to every $v \in \operatorname{coder} f(\bar{x})$ (but not relative to any $v \notin \operatorname{coder} f(\bar{x})$). Moreover, the subgradient mapping $\operatorname{coder} f$ *is then proto-differentiable at $\bar{x}$ relative to every $v \in \operatorname{coder} f(\bar{x})$, with*

$$(\operatorname{coder} f)'_{\bar{x},v}(x') = \operatorname{coder} \big(\tfrac{1}{2} f''_{\bar{x},v} \big)(x') \quad \text{for all } x'. \tag{3.2}$$

Formula (3.2), relating the proto-derivative of the subgradient mapping to the subgradients of the second-order epi-derivative was first established in the case of general convex functions f in Rockafellar [17], under the assumption that either the proto-derivative or the second-order epi-derivative exists. This formula was subsequently extended to the setting of Theorem 3.1 by Poliquin [21], and recently, by Poliquin [22] even further to the composition of an arbitrary lower semicontinuous convex function and a $\mathcal{C}^2$ mapping satisfying the constraint qualification (1.1). An infinite-dimensional version of the original result in the convex case has been obtained by Do [20]. There are deep applications to the study of perturbations of optimal solutions and associated multipliers in parametric optimization; some of these applications will be discussed in section 4.

An immediate consequence of the first-order formula in Theorem 3.1 is that in the case of a fully amenable function f, the set of subgradients of f at $\bar{x}$ is a *polyhedral* set, and the epi-derivative f'_x is a *piecewise linear, positively homogeneous (of degree 1) convex function.*

In [7] we proved further that: *the second-order epi-derivative of a fully amenable function is the sum of a piecewise linear-quadratic convex function homogeneous of degree 2 and a quadratic function.* By using formula (3.2) we obtain the following subgradient version: *the proto-derivative of the subgradient mapping of a fully amenable function is* *the sum of a polyhedral (in the sense of Robinson) homogeneous piecewise linear maximal monotone set-valued mapping and a symmetric linear transformation.*

We now excerpt the calculus results from [7] that will be needed in our study of problem $(\mathcal{P})$.

Theorem 3.2 (addition rule) [7, Theorem 3.1]. *Assume the functions $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}$ for $i = 1, 2$ are amenable at $\bar{x}$ and such that*

$$\text{if } v_1 + v_2 = 0 \text{ with } v_i \in N(\bar{x}|\,\mathrm{dom}\, f_i), \text{ then } v_1 = v_2 = 0. \tag{3.4}$$

Then the function $f = f_1 + f_2$ is amenable at all points x in some neighborhood of $\bar{x}$ relative to $\mathrm{dom}\, f = \mathrm{dom}\, f_1 \cap \mathrm{dom}\, f_2$, with

$$\begin{aligned} f'_x(x') &= (f'_1)_x(x') + (f'_2)_x(x'), \\ \partial f(x) &= \partial f_1(x) + \partial f_2(x), \\ N(x|\,\mathrm{dom}\, f) &= N(x|\,\mathrm{dom}\, f_1) + N(x|\,\mathrm{dom}\, f_2). \end{aligned} \tag{3.5}$$

If each f_i is fully amenable at $\bar{x}$, there is the additional conclusion that f is fully amenable at such neighboring points x, with

$$f''_{x,v}(x') = \max_{\substack{v_1+v_2=v \\ v_i \in \partial f_i(x)}} \Big\{ (f_1)''_{x,v_1}(x') + (f_2)''_{x,v_2}(x') \Big\} \text{ for all } v \in \partial f(x), \tag{3.6}$$

and in terms of the set $V_v(x, x')$ giving the elements (v_1, v_2) for which the maximum in this formula is achieved, also

$$(\mathrm{coder}\, f)'_{x,v}(x') = \bigcup_{(v_1,v_2)\in V_v(x,x')} \Big\{ (\mathrm{coder}\, f_1)'_{x,v_1}(x') + (\mathrm{coder}\, f_2)'_{x,v_2}(x') \Big\} \tag{3.7}$$

Condition (3.4) is trivially satisfied, for instance, when $\bar{x}$ belongs to the interior of either $\mathrm{dom}\, f_1$ or $\mathrm{dom}\, f_2$.

Theorem 3.3 (chain rule) [7, Theorem 3.5]. *Suppose $f(x) = g\big(F(x)\big)$ for a C^1 mapping $F : \mathbb{R}^n \to \mathbb{R}^d$ and a function $g : \mathbb{R}^d \to \overline{\mathbb{R}}$. Let $\bar{x}$ be a point such that g is amenable at $F(\bar{x})$ and*

$$\text{there is no } y \neq 0 \text{ in } N\big(F(\bar{x})|\,\mathrm{dom}\, g\big) \text{ with } \nabla F(\bar{x})^* y = 0. \tag{3.8}$$

Then f is amenable at all points x in some neighborhood of $\bar{x}$ relative to $\mathrm{dom}\, f$, with

$$\begin{aligned} \partial f(x) &= \nabla F(x)^* \partial g\big(F(x)\big), \\ N(x|\,\mathrm{dom}\, f) &= \nabla F(x)^* N\big(F(x)|\,\mathrm{dom}\, g\big), \\ f'_x(x') &= g'_{F(x)}\big(\nabla F(x) x'\big). \end{aligned}$$

If g is fully amenable at $F(\bar{x})$ and F is a $\mathcal{C}^2$ mapping, then f is fully amenable at all such neighboring points x, with

$$\begin{aligned} f''_{x,v}(x') = \max_{y \in Y_v(x)} \Big\{ g''_{F(x),y}\big(\nabla F(x)x'\big) + \Big\langle x', \nabla^2 \langle y, F\rangle(x)x' \Big\rangle \Big\} \\ \text{where } Y_v(x) = \Big\{ y \in \partial g(F(x)) \,\Big|\, \nabla F(x)^* y = v \Big\}, \end{aligned} \tag{3.9}$$

and in terms of the set $Y'_v(x,x')$ of vectors y achieving the maximum in this formula, also

$$(\text{coder}\, f)'_{x,v}(x') = \bigcup_{y \in Y'_v(x,x')} \Big\{ \nabla F(x)^* (\text{coder}\, g)'_{F(x),y}\big(\nabla F(x)x'\big) + \nabla^2 \langle y, F\rangle(x)x' \Big\}. \tag{3.10}$$

Here the function $\langle y, F\rangle : I\!R^n \to I\!R$ is given by $\langle y, F\rangle(x) := \langle y, F(x)\rangle$.

As an immediate application of these rules to problem $(\mathcal{P})$ in Section 1, we prove the following facts about the essential objective function in that problem.

Theorem 3.4. *In the amenable case of problem $(\mathcal{P})$, let $\bar{x}$ be a feasible solution at which the basic constraint qualification in Definition 1.3 is satisfied. Then the essential objective function $f = k + h \circ G$ in $(\mathcal{P})$ is amenable at $\bar{x}$, and*

$$\begin{aligned} f'_{\bar{x}}(x') &= k'_{\bar{x}}(x') + h'_{G(\bar{x})}\big(\nabla G(\bar{x})x'\big), \\ \partial f(\bar{x}) &= \partial k(\bar{x}) + \nabla G(\bar{x})^* \partial h\big(G(\bar{x})\big). \end{aligned} \tag{3.11}$$

In the fully amenable case, f is fully amenable at $\bar{x}$, and one has in addition that, for all $v \in \partial f(\bar{x})$,

$$\begin{aligned} f''_{\bar{x},v}(x') = \max_{\substack{y \in Y_v(\bar{x}) \\ w = v - \nabla G(\bar{x})^* y}} \Big\{ k''_{\bar{x},w}(x') + h''_{G(\bar{x}),y}\big(\nabla G(\bar{x})x'\big) + \big\langle x', \nabla^2 \langle y, G\rangle(\bar{x})x' \big\rangle \Big\} \\ \text{with } Y_v(\bar{x}) = \Big\{ y \in \partial h\big(G(\bar{x})\big) \,\Big|\, v - \nabla G(\bar{x})^* y \in \partial k(\bar{x}) \Big\}, \end{aligned} \tag{3.12}$$

and in terms of the set $Y'_v(\bar{x},x')$ consisting of the vectors y achieving the maximum in this formula, also

$$\begin{aligned} &(\text{coder}\, f)'_{\bar{x},v}(x') = \\ &\bigcup_{\substack{y \in Y'_v(\bar{x},x') \\ w = v - \nabla G(\bar{x})^* y}} \Big\{ (\text{coder}\, k)'_{\bar{x},w}(x') + \nabla G(\bar{x})^* (\text{coder}\, h)'_{G(\bar{x}),y}\big(\nabla G(\bar{x})x'\big) + \nabla^2 \langle y, G\rangle(\bar{x})x' \Big\}. \end{aligned} \tag{3.13}$$

Proof. The basic constraint qualification in Definition 1.3 implies in particular that there is no $y \neq 0$ in $N\big(G(\bar{x}) \big| \operatorname{dom} h\big)$ with $\nabla G(\bar{x})^* y = 0$ (because $0 \in N(\bar{x} | \operatorname{dom} h)$). This observation enables us to apply Theorem 3.3 to the function $h\big(G(x)\big)$, from which we get in particular

$$N(\bar{x}|\,\mathrm{dom}\, h\circ G) = \nabla G(\bar{x})^* N\big(G(\bar{x})\big|\,\mathrm{dom}\, h\big).$$

The theorem now follows by invoking the results of Theorem 3.3 together with those of Theorem 3.2 (taking $f_1 = k$ and $f_2 = h\circ G$; in this situation condition (1.4) is equivalent to condition (3.4)). □

We obtain optimality conditions now through the principles stated in Section 2.

Theorem 3.5. *In the fully amenable case of problem $(\mathcal{P})$, consider a feasible solution $\bar{x}$ along with the multiplier set*

$$Y(\bar{x}) := \Big\{\bar{y} \in \mathrm{coder}\, h\big(G(\bar{x})\big) \,\Big|\, -\nabla G(\bar{x})^*\bar{y} \in \mathrm{coder}\, k(\bar{x})\Big\}. \tag{3.14}$$

(a) (necessary conditions): If $\bar{x}$ is locally optimal and satisfies the basic constraint qualification in Definition 1.3, then $Y(\bar{x}) \neq \emptyset$, and for every x' one has

$$k''_{\bar{x},\bar{w}}(x') + h''_{G(\bar{x}),\bar{y}}\big(\nabla G(\bar{x})x'\big) + \Big\langle x', \nabla^2\langle \bar{y}, G\rangle(\bar{x})x'\Big\rangle \geq 0$$
$$\text{for some } \bar{y} \in Y(\bar{x}), \text{ where } \bar{w} = -\nabla G(\bar{x})^*\bar{y}.$$

(b) (sufficient conditions): If $\bar{x}$ is such that $Y_0(\bar{x}) \neq \emptyset$, and for every $x' \neq 0$ one has

$$k''_{\bar{x},\bar{w}}(x') + h''_{G(\bar{x}),\bar{y}}\big(\nabla G(\bar{x})x'\big) + \Big\langle x', \nabla^2\langle \bar{y}, G\rangle(\bar{x})x'\Big\rangle > 0$$
$$\text{for some } \bar{y} \in Y(\bar{x}), \text{ where } \bar{w} = -\nabla G(\bar{x})^*\bar{y},$$

then $\bar{x}$ is locally optimal in the strong sense.

Example 3.6. In the case of problem $(\mathcal{P}_0)$ in Example 1.4, the conditions in Theorem 3.5 give a close generalization of well known optimality conditions in nonlinear programming. First of all, because $k = f_0 + \delta_C$ we have $\mathrm{coder}\, k(\bar{x}) = \nabla f_0(\bar{x}) + N(\bar{x}|C)$. Likewise $\mathrm{coder}\, h\big(G(\bar{x})\big) = N\big(G(\bar{x})\big|D\big)$, where the latter is the product of the intervals $N\big(f_i(\bar{x})\big|I_i\big)$. The multiplier set in (3.14) thus comes out as

$$Y(\bar{x}) = \Big\{\bar{y} \in I\!R^m \,\Big|\, \bar{y}_i \in N\big(f_i(\bar{x})\big|I_i\big) \text{ for } i \in [1,m],\ -\nabla_x L(\bar{x},\bar{y}) \in N(\bar{x}|C)\Big\} \tag{3.15}$$

for the Lagrangian

$$L(x,y) := f_0(x) + y_1 f_1(x) + \cdots + y_m f_m(x). \tag{3.16}$$

The first-order condition just asserts for the given $\bar{x}$ the existence of some $\bar{y}$ in this set. To see what the second-order condition amounts to, we need to work out the form of the second-order epi-derivatives that are involved. Application of Theorem 3.2 to k gives us

$$k''_{\bar{x},\bar{w}}(x') = \big\langle x', \nabla^2 f_0(\bar{x})x'\big\rangle + (\delta_C)''_{\bar{x},s}(x') \text{ for } s = \bar{w} - \nabla f_0(\bar{x}).$$

On the other hand, because C is polyhedral, we have for any $s \in \operatorname{coder}\delta_C(\bar{x}) = N(\bar{x}|C)$ that

$$(\delta_C)''_{\bar{x},s}(x') = \begin{cases} 0 & \text{if } x' \in T(\bar{x}|C),\ \langle s, x'\rangle = 0, \\ \infty & \text{otherwise,} \end{cases}$$

where $T(\bar{x}|C)$ is the tangent cone to C at $\bar{x}$ in the sense of convex analysis; this is the polar of the normal cone $N(\bar{x}|C)$. The polarity implies in the case of $s = \bar{w} - \nabla f_0(\bar{x})$ that

$$\langle f_0(\bar{x}), x'\rangle \geq \langle \bar{w}, x'\rangle \text{ for all } x' \in T(\bar{x}|C). \tag{3.17}$$

In the optimality conditions in question, $\bar{w}$ stands for the vector $-\nabla G(\bar{x})^*\bar{y}$.

Similarly we have for general $\bar{z} \in D$, $z' \in I\!R^m$, and $\bar{y} \in \operatorname{coder} h(\bar{z}) = N\big(G(\bar{x})\big|D\big)$ that

$$h''_{\bar{z},\bar{y}}(z') = (\delta_D)''_{\bar{z},\bar{y}}(z') = \begin{cases} 0 & \text{if } z' \in T(\bar{z}|D),\ \langle \bar{y}, z'\rangle = 0, \\ \infty & \text{otherwise.} \end{cases}$$

We are interested in this formula in the case of $\bar{z} = G(\bar{x})$ and $z' = \nabla G(\bar{x})x' \in T(\bar{z}|D)$, and then $\langle \bar{y}, z'\rangle = \langle \nabla G(\bar{x})^*\bar{y}, x'\rangle = -\langle \bar{w}, x'\rangle$, so that

$$\langle \bar{w}, x'\rangle \geq 0 \text{ for all } x' \text{ with } \nabla G(\bar{x})x' \in T\big(G(\bar{x})\big|D\big). \tag{3.18}$$

Summarizing so far, we see that the second-order expression in the optimality conditions in Theorem 3.5 has for any given x' the value

$$\begin{cases} \big\langle x', \big[\nabla^2 f_0(\bar{x}) + \nabla^2\langle \bar{y}, G\rangle(\bar{x})\big]x'\big\rangle & \text{if } x' \in T(\bar{x}|C),\ \langle f_0(\bar{x}), x'\rangle = \langle \bar{w}, x'\rangle, \\ & \quad \nabla G(\bar{x})x' \in T(\bar{z}|D),\ \langle \bar{w}, x'\rangle = 0, \\ \infty & \text{otherwise} \end{cases}$$
$$= \begin{cases} \langle x', \nabla^2_{xx} L(\bar{x}, \bar{y})x'\rangle & \text{if } x' \in T(\bar{x}|C),\ \langle \nabla f_0(\bar{x}), x'\rangle = 0, \\ & \quad \nabla G(\bar{x})x' \in T(\bar{z}|D), \\ \infty & \text{otherwise,} \end{cases}$$

where the simplified condition on the right results from (3.17) and (3.18). The tangent cone $T\big(G(\bar{x})\big|D\big)$ is the product of the intervals $T\big(f_i(\bar{x})\big|I_i\big)$, whereas

$$\nabla G(\bar{x})x' = \Big(\big\langle \nabla f_1(\bar{x}), x'\big\rangle, \ldots, (\big\langle \nabla f_m(\bar{x}), x'\big\rangle\Big). \tag{3.19}$$

We have arrived at the following optimality conditions for problem $(\mathcal{P}_0)$. in terms of the multiplier set in (3.15), the Lagrangian function (3.16), and the polyhedral cone

$$X(\bar{x}) = \Big\{ x' \in T(\bar{x}|C) \;\Big|\; \langle \nabla f_0(\bar{x}), x' \rangle = 0, \quad \langle \nabla f_i(\bar{x}), x' \rangle \in T\big(f_i(\bar{x}) \big| I_i\big) \text{ for } i \in [1,m] \Big\}. \tag{3.20}$$

Necessary conditions. If $\bar{x}$ is locally optimal in (P_0) and satisfies the basic constraint qualification in (1.6), then $Y(\bar{x}) \neq \emptyset$, and for every $x' \in X(\bar{x})$ there exists $\bar{y} \in Y(\bar{x})$ such that $\langle x', \nabla^2_{xx} L(\bar{x}, \bar{y}) x' \rangle \geq 0$.

Sufficient conditions. If $\bar{x}$ is such that $Y(\bar{x}) \neq \emptyset$, and for every nonzero $x' \in X(\bar{x})$ there exists $\bar{y} \in Y(\bar{x})$ such that $\langle x', \nabla^2_{xx} L(\bar{x}, \bar{y}) x' \rangle > 0$, then $\bar{x}$ is locally optimal in the strong sense.

Note that the sets $N\big(f_i(\bar{x}) \big| I_i\big)$ and $T\big(f_i(\bar{x}) \big| I_i\big)$ in definitions (3.15) and (3.20) are closed intervals which must be of one of the four types $(-\infty, 0]$, $[0, \infty)$, $(-\infty, \infty)$ or $[0, 0]$, depending only on the position of the value $f_i(\bar{x})$ in the interval I_i.

4. APPLICATIONS TO PERTURBATION OF SOLUTIONS

A fundamental issue in optimization theory is how the optimal value and the optimal solution set for a given problem may be affected by shifts in the data parameters on which the problem depends. There is a long history to the study of perturbations of the optimal value. This has been a key to understanding duality in convex programming, and it is also recognized now as one of the best ways of deriving and interpreting Lagrange multiplier vectors, even in problems without convexity. The study of perturbations of optimal solutions has until recently been pursued in a much more limited way. Most of the effort has been dedicated to the important task of determining when the optimal solution is unique and, along with perhaps the corresponding multiplier vector, exhibits Lipschitz dependence on the parameters; here the work of Robinson [26] [27] [28] [29] has set the pace. Results of Aubin [30], which appeared to offer wider generality in a setting of parametric convex optimization, hinged on assumptions that turned out to be satisfiable only when the solution and multiplier vector were unique after all and depended in a strictly differentiable way on the parameters, cf. Rockafellar [31].

Recently, however, interest has picked up in trying better to understand what happens under less stringent assumptions, where a nontrivial *set* of optimal solutions could exist: Shapiro [32], Rockafellar [10] [18], Kyparisis [33], King and Rockafellar [34] [35].

As an application of the calculus results for amenable functions, we consider now a parameterized class of problems having the form of $(\mathcal{P})$:

$$\big(\mathcal{P}(u,v)\big) \qquad \text{minimize } \; k(x)+h\big(G(x)+u\big)+\langle x,v\rangle \; \text{ over all } \; x\in I\!R^n.$$

The parameter vectors $u\in I\!R^m$ and $v\in I\!R^n$ give the basic primal and dual perturbations relative to problem $(\mathcal{P})$, which corresponds to $(u,v)=(0,0)$. Our attention will be directed to the fully amenable case, where k and h are fully amenable and G is $\mathcal{C}^2$. Instead of analyzing perturbations of optimal solutions directly, we shall analyze perturbations of pairs (x,y) satisfying the first-order optimality conditions for $\big(P(u,v)\big)$ in the following sense.

Proposition 4.1. *If $\bar{x}$ is locally optimal for problem $\big(\mathcal{P}(\bar{u},\bar{v})\big)$ and*

$$\textit{there is no } y\neq 0 \textit{ in } N\big(G(\bar{x})+\bar{u}\,\big|\,\operatorname{dom} h\big) \textit{ with } -\nabla G(\bar{x})^*y\in N(\bar{x}|\operatorname{dom} k), \tag{4.1}$$

then there exists a vector $\bar{y}$ such that

$$\bar{y}\in\partial h\big(G(\bar{x})+\bar{u}\big), \qquad -\nabla G(\bar{x})^*\bar{y}-\bar{v}\in\operatorname{coder} k(\bar{x}). \tag{4.2}$$

In the convex case, where k and h are convex, and so is $\langle y,G\rangle$ for all $y\in\operatorname{dom} h^*$ (the conjugate of h), this condition is in fact sufficient for $\bar{x}$ to be globally optimal.

Proof. The necessary condition is an immediate specialization of the first-order part of Theorem 3.5(a), obtained by replacing $k(x)$ by $\widehat{k}(x)=k(x)+\langle\bar{v},x\rangle$ and $h(w)$ by $\widehat{h}(w)=h(w+\bar{u})$. The sufficiency in the convex case is a fact of convex analysis; the convexity condition implies the convexity of the essential objective function in the problem. □

Our object of study will be the mapping $M: I\!R^m\times I\!R^n\to I\!R^n\times I\!R^m$ defined by

$$M(u,v)=\Big\{(x,y)\,\Big|\, y\in\operatorname{coder} h\big(G(x)+u\big),\; -\nabla G(x)^*y-v\in\operatorname{coder} k(x)\Big\}, \tag{4.3}$$

which associates with the parameter vectors u and v the vectors x and y appearing in the necessary condition in Proposition 4.1. In the convex case, the pairs (x,y) in $M(u,v)$ are the saddle points of a certain Lagrangian function, so that not only is x optimal in $\big(\mathcal{P}(u,v)\big)$, but y is optimal in a certain dual problem $\big(\mathcal{D}(u,v)\big)$. In the nonconvex case, optimality cannot be claimed in general, since only first-order conditions are involved. Nonetheless, the mapping M is attractive for technical reasons, because results can be obtained in situations where, if one tried to concentrate only on optimal solutions, relatively little could be said without far stronger assumptions.

Theorem 4.2. *Suppose $(\bar{x},\bar{y}) \in M(\bar{u},\bar{v})$. Without any additional assumptions, M is proto-differentiable at $(\bar{u},\bar{v})$ relative to $(\bar{x},\bar{y})$, and the corresponding proto-derivative mapping is given in terms of $\bar{w} = -\bar{v} - \nabla G(\bar{x})\bar{y}$ by*

$$(x',y') \in M'_{(\bar{u},\bar{v}),(\bar{x},\bar{y})}(u',v') \iff \left\{ \begin{array}{l} y' \in (\operatorname{coder} h)'_{G(\bar{x})+\bar{u},\,\bar{y}}\big(\nabla G(\bar{x})x' + u'\big), \\ -v' - \nabla G(\bar{x})^* y' - \nabla^2\langle \bar{y}, G\rangle(\bar{x})x' \in \operatorname{coder} k'_{\bar{x},\bar{w}}(x'). \end{array} \right\} \tag{4.4}$$

Proof. Consider the function $\varphi(x,u) := k(x) + h\big(G(x)+u\big)$. This can be viewed as $\psi\circ H$ for $\psi(x,z) = k(x) + h(z)$ and $H(u,x) = \big(x, G(x)+u\big)$. Clearly H is again $\mathcal{C}^2$, and $\psi(x,z)$ is fully amenable (everywhere in its effective domain); this can be surmised from the definition of full amenability, but it can also be verified formally from Theorem 3.2. Because H has nonsingular Jacobian, we can apply the chain rule in Theorem 3.3 to calculate the various derivatives of φ. In this way we obtain in particular that in a neighborhood of $\bar{x},\bar{u}$,

$$\operatorname{coder}\varphi(x,u) = \Big\{(-v,y)\,\Big|\, y \in \operatorname{coder} h\big(G(x)+u\big),\ -v - \nabla G(x)^* y \in \operatorname{coder} k(x)\Big\}, \tag{4.5}$$

and furthermore in terms of $\bar{w} = -\bar{v} - \nabla G(\bar{x})^*\bar{y}$ that

$$\operatorname{coder}\varphi'_{(\bar{x},\bar{y}),(-\bar{v},\bar{y})}(x',u') = \Big\{(-v',y')\,\Big|\, y' \in (\operatorname{coder} h)'_{\left(G(\bar{x})+\bar{u},\,\bar{y}\right)}\big(\nabla G(\bar{x})x' + u'\big), \\ -v' - \nabla G(\bar{x})^* y' - \nabla^2\langle \bar{y}, G\rangle(\bar{x})x' \in \operatorname{coder} k'_{\bar{x},\bar{w}}(x')\Big\}. \tag{4.6}$$

Formula (4.5) implies

$$M(u,v) = \Big\{(x,y)\,\Big|\,(-v,y) \in \partial\varphi(x,u)\Big\}. \tag{4.7}$$

Inasmuch as proto-differentiation is a geometric property of graphs which is preserved under permutations of arguments, we deduce from (4.7) that

$$(x',y') \in M'_{(\bar{u},\bar{v}),(\bar{x},\bar{y})}(u',v') \iff (-v',y') \in \big(\operatorname{coder}\varphi\big)'_{(\bar{x},\bar{u}),(-\bar{v},\bar{y})}(x',u'). \tag{4.8}$$

The latter condition means by (4.6) precisely the relations on the right in (4.4). □

The proto-derivative formula in Theorem 4.2 involves conditions of the same character as the ones defining M in (4.3). We are led in this way to introduce as an auxiliary to the given problem $\big(\mathcal{P}(\bar{u},\bar{v})\big)$ and the solution vectors $\bar{x}$ and $\bar{y}$ the following *derivative problem*:

$\big(\mathcal{P}'_{(\bar{u},\bar{v}),(\bar{x},\bar{y})}(u',v')\big)$ minimize $\bar{k}(x') + \bar{h}\big(\bar{G}(x') + u'\big) + \langle v', x'\rangle$ in x', where one defines (again with $\bar{w} = -\bar{v} - \nabla G(\bar{x})^* \bar{y}$):

$$\begin{aligned} \bar{k}(x') &= \tfrac{1}{2} k''_{\bar{x},\bar{w}}(x') + \tfrac{1}{2}\Big\langle x', \nabla^2 \langle \bar{y}, G\rangle(\bar{x}) x' \Big\rangle, \\ \bar{h}(z') &= \tfrac{1}{2} h''_{G(\bar{x})+\bar{u},\,\bar{y}}(z'), \qquad \bar{G}(x') = \nabla G(\bar{x}) x'. \end{aligned} \tag{4.9}$$

We then obtain an interpretation of Theorem 4.2 as well as a prescription for actually calculating the perturbation vectors in any given case. *The pairs* $(x', y') \in M'_{(\bar{u},\bar{v}),(\bar{x},\bar{y})}(u', v')$ *are precisely the ones satisfying the first-order conditions for problem* $\big(\mathcal{P}'_{(\bar{u},\bar{v}),(\bar{x},\bar{y})}(u', v')\big)$ (as singled out by Proposition 4.1, with the elements in (4.9) replacing the original ones).

Example 4.3. Specializing to the case of problem $(\mathcal{P}_0)$ in Example 1.4, we look at the parameterized problem

$$\big(\mathcal{P}_0(u, v)\big) \qquad \begin{aligned} &\text{minimize } f_0(x) + \langle v, x\rangle \text{ over all } x \in C \text{ satisfying} \\ &\qquad f_i(x) + u_i \in I_i \text{ for } i = 1, \ldots, m. \end{aligned}$$

The multiplier set in parametric terms is

$$\begin{aligned} Y(\bar{u}, \bar{v}, \bar{x}) = \Big\{ \bar{y} \in I\!R^m \;\Big|\; &\bar{y}_i \in N\big(f_i(\bar{x}) + u_i \big| I_i\big) \text{ for } i = 1, \ldots, m, \\ &- \nabla_x L(\bar{x}, \bar{y}) - \bar{v} \in N(\bar{x}|C) \Big\}, \end{aligned} \tag{4.10}$$

where L is the Lagrangian in (3.16). The second-derivative calculations in Example 3.6 reveal that the corresponding derivative problem is:

$$\big(\mathcal{P}_0{}'_{(\bar{u},\bar{v}),(\bar{x},\bar{y})}(u', v')\big) \qquad \begin{aligned} &\text{minimize } \tfrac{1}{2}\big\langle x', \nabla^2_{xx} L(\bar{x}, \bar{y}) x' \big\rangle + \big\langle v', x' \big\rangle \\ &\text{over all } x' \in T(\bar{x}|C) \text{ satisfying } \big\langle \nabla f_0(\bar{x}), x' \big\rangle = 0, \\ &\big\langle \nabla f_i(\bar{x}), x' \big\rangle + u'_i \in T\big(f_i(\bar{x}) \big| I_i\big) \text{ for } i = 1, \ldots, m. \end{aligned}$$

Thus, in reference to the given choice of $(\bar{u}, \bar{v})$ and the pair $(\bar{x}, \bar{y}) \in M(\bar{x}, \bar{y})$ (not necessarily unique) for the given problem, if we select a pair (u', v') giving a proposed "rate of change" of $(\bar{u}, \bar{v})$ and wish to determine all the pairs (x', y') realizable as a "rate of change" of $(\bar{x}, \bar{y})$, we can do so by calculating all the pairs (x', y') satisfying the first-order optimality condition (generalized Kuhn-Tucker pairs) for this auxiliary problem.

In the traditional case of $\big(\mathcal{P}_0(\bar{u}, \bar{v})\big)$ in which the set C is all of $I\!R^n$ and each interval I_i is either $(-\infty, 0]$ or $[0, 0]$, the derivative problem has an even more special form, which has appeared extensively in the literature; see for example Kyparisis [33]. Our analysis sheds a strong new light on the the nature of this auxiliary problem.

REFERENCES:

1. J. V . Burke, "An exact penalization viewpoint of constrained optimization," *SIAM Journal on Control and Optimization*, **29** (1991), 968–998.
2. J. V. Burke and R. A. Poliquin, "Optimality conditions for non- finite valued convex composite functions," *Math. Programming*, to appear.
3. R. Fletcher, "A model algorithm for composite nondifferentiable optimization problems," *Math. Programming Study 17* (1982) 67–76.
4. A. D. Ioffe, "Variational analysis of a composite function: a formula for the lower second order directional derivative," *J. Math. Anal. Appl.*, forthcoming.
5. R. T. Rockafellar, "First- and second-order epi-differentiability in nonlinear programming," *Trans. Amer. Math. Soc.* **307** (1988), 75–107.
6. R. T. Rockafellar, "Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives," *Math. of Oper. Research* **14** (1989), 462–484.
7. R. A. Poliquin and R. T. Rockafellar, "A calculus of epi-derivatives applicable to optimization," preprint.
8. S. M. Robinson, "Some continuity properties of polyhedral multifunctions," *Math. Programming Study* **14** (1981), 206–214.
9. J. Sun, *On Monotropic Piecewise Quadratic Programming*, dissertation, Department of Applied Math., University of Washington, August, 1986.
10. 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.
11. B. S. Mordukhovich, *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow, 1988 (in Russian; English translation to appear).
12. F. H. Clarke, "Generalized gradients and applications," *Transactions Amer. Math. Soc.* **205** (1975), 247–262.
13. F. H. Clarke, *Optimization and Nonsmooth analysis*, Wiley, 1983.
14. R. J. -B. Wets,"Convergence of convex functions, variational inequalities and convex optimization problems," in *Variational Inequalities and Complementary Problems*, R. Cottle, F. Giannessi and J.-L. Lions (eds.), John-Wiley and Sons.
15. H. Attouch, *Variational Convergence of Functions and Operators*, Pitman, London, 1984.

16. H. Attouch and R. J. -B. Wets, "Epigraphical analysis," *Annales de l'Institût Henri Poincaré, Analyse Non-Linéaire*, **6** (1989), 73–100.

17. R. T. Rockafellar, "Generalized second derivatives of convex functions and saddle functions," *Transactions Amer. Math. Soc.* **320** (1990), 810–822.

18. 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.

19. R. Cominetti, "On Pseudo-differentiability," *Transactions Amer. Math. Soc.*, to appear.

20. C. N. Do, "Generalized second derivatives of convex functions in reflexive Banach spaces," *Transactions Amer. Math. Soc.*, to appear.

21. R. A. Poliquin, "Proto-differentiation of subgradient set-valued mappings," *Canadian Math J.* **42** (1990), 520–532.

22. R. A. Poliquin, "An extension of Attouch's Theorem and its application to second-order epi-differentiation of convex composite functions," preprint.

23. R. A. Poliquin and R. T. Rockafellar, "Differentiability of solution mappings and sensitivity analysis in parametric optimization," preprint.

24. I. D. Ioffe, "Calculus of Dini subdifferentials of functions and contingent coderivatives of set-valued maps," *Nonlinear Analysis*, **8** (1984), 517–539.

25. 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.

26. S. M. Robinson, "Generalized equations and their solutions, part I: basic theory," *Math. Programming Study* **10** (1979), 128–141.

27. S. M. Robinson, "Strongly regular generalized equations," *Math. Oper. Research* **5** (1980), 43–62.

28. S. M. Robinson, "Implicit B-differentiability in generalized equations," MRC Technical Summary Report #2854, Univ. of Wisconsin (1985).

29. S. M. Robinson, "An implicit-function theorem for a class of nonsmooth functions," *Math. Oper. Research* **16** (1991), 292–309.

30. J.-P. Aubin, "Lipschitz behavior of solutions to convex minimization,"*Math. Oper. Res.* **9** (1984), 87–111. **22** (1976), 130–185.

31. R. T. Rockafellar, "Maximal monotone relations and the second derivatives of nonsmooth functions," *Ann. Institût H. Poincaré: Analyse Non Linéaire* **2** (1985), 167–184.

32. A. Shapiro, "Perturbation theory of nonlinear programs when the set of optimal solutions is not a singleton," *Appl. Math. Opt.* **18** (1988),215–229.

33. J. Kyparisis, "Sensitivity analysis for nonlinear programs and variational inequalities with nonunique multipliers," *Math. of Oper. Research* **15** (1990), 286–298.

34. A. J. King and R. T. Rockafellar, "Sensitivity analysis for nonsmooth generalized equations," *Transactions Amer. Math. Soc.*, to appear.

35. A. J. King and R. T. Rockafellar, "Asymptotic theory for solutions in generalized M-estimation and stochastic programming," *Math. Oper. Research*, to appear.