VARIATIONAL CONDITIONS AND THE PROTO-DIFFERENTIATION OF PARTIAL SUBGRADIENT MAPPINGS

A. B. Levy and R. T. Rockafellar*

Abstract. Subgradient mappings have many roles in variational analysis, such as the formulation of optimality conditions and, in the special case of normal cone mappings, the statement and analysis of variational inequalities and related expressions. Central to the study of perturbations of solutions to systems of conditions in which subgradient mappings appear are concepts of generalized differentiation such as proto-differentiability, which is unhampered by a solution mapping's potential multivaluedness.

This paper extends the known examples of proto-differentiability by showing that a large and important class of "partial" subgradient mappings have this property. Until now the main examples have been the complete subgradient mappings associated with fully amenable functions. The extension, made difficult by the need to deal geometrically with projections of graphs, relies on the notion of a fully amenable function having additional variables that provide a "compatible parameterization." The results are applied to the sensitivity analysis of generalized variational inequalities in which the underlying set need not be convex and can vary with the parameters.

Keywords. Variational analysis, nonsmooth analysis, sensitivity analysis, variational inequalities, generalized equations, optimization, proto-derivatives, amenable functions.

1991 *Mathematics Subject Classification.* Primary 58C15; Secondary 58C20

November, 1994

* This work was supported in part by the National Science Foundation under grant DMS–9200303 and the U.S.-Israel Binational Science Foundation under grant 90–00455/2.

1. Introduction

Associated with any lower semicontinuous, proper function $f : I\!R^n \to \overline{I\!R} := I\!R \cup \{\pm\infty\}$ is the multifunction $\partial f : x \in I\!R^n \rightrightarrows v \in I\!R^n$ giving the set of subgradient vectors v of f at a point x (taken here in the sense of limiting proximal subgradients). This set-valued subgradient mapping has many uses in nonsmooth analysis, particularly in connection with optimality rules and their parameterization. Indeed, the relation $\bar{v} \in \partial f(\bar{x})$ is itself a necessary condition for $\bar{x}$ to give a local minimum in the problem of minimizing $f(x) - \langle \bar{v}, x\rangle$ with respect to x. The inverse multifunction $(\partial f)^{-1}$, which assigns to v the set of all x satisfying $v \in \partial f(x)$, thus depicts, in its parametric dependence on v, a set of "quasi-solutions" to the problem of minimizing $f(x) - \langle v, x\rangle$ in x. An important question in this and many similar situations is how to get a handle on possible "rates of change" of solutions with respect to perturbations of parameters despite the set-valuedness, or even if the sets are singletons, the typical lack of smoothness in such dependence.

A concept of differentiation that has turned out to be fruitful in analyzing multifunctions is proto-differentiability, which was introduced in [.optimi rockaf proto-d set-val.]. Consider any multifunction $\Gamma : I\!R^d \rightrightarrows I\!R^n$ and any pair $(\bar{w}, \bar{z})$ in the graph of Γ, i.e., with $\bar{z} \in \Gamma(\bar{w})$. For each $t > 0$ one can form the difference quotient multifunction

$$(\Delta_t\Gamma)_{\bar{w},\bar{z}} : \omega \mapsto \big[\Gamma(\bar{w} + t\omega) - \bar{z}\big]/t. \tag{1.1}$$

If, as $t \searrow 0$, the multifunctions $(\Delta_t\Gamma)_{\bar{w},\bar{z}}$ converge graphically to a multifunction D, in the sense that their graphs converge to that of D as subsets of $I\!R^d \times I\!R^n$ (as will be explained more fully in Section 2), Γ is said to be *proto-differentiable at $\bar{w}$ for $\bar{z}$.* The limit D is the *proto-derivative* multifunction and is denoted by $\Gamma'_{\bar{w},\bar{z}}$. It associates with each $\omega \in I\!R^d$ a subset $\Gamma'_{\bar{w},\bar{z}}(\omega) \subset I\!R^n$, which for some choices of ω could be empty.

Proto-differentiability of a multifunction amounts to a tangent-cone property of its graph, moreover one which is preserved under linear or even smooth nonlinear transformations of the graph. For example, Γ is proto-differentiable at $\bar{w}$ for $\bar{z}$ if and only if Γ^{-1} is proto-differentiable at $\bar{z}$ for $\bar{w}$, and then the corresponding proto-derivative multifunctions are themselves inverses of each other: $(\Gamma^{-1})'_{\bar{z},\bar{w}} = (\Gamma'_{\bar{w},\bar{z}})^{-1}$. But the price paid for this accommodating geometry is the need for coping with a separate proto-derivative multifunction $\Gamma'_{\bar{w},\bar{z}}$ at $\bar{w}$ for each choice of $\bar{z} \in \Gamma(\bar{w})$.

Much effort has been devoted to identifying circumstances in which a multifunction Γ is proto-differentiable, and if so, determining useful formulas for the proto-derivatives. For the case of a subgradient mapping ∂f, a key result was obtained by Poliquin [.poliquin set-val can subgrad proto-d¡, Theorem 4.6¿.] in drawing on a generalized second-derivative theory of Rockafellar [.rockaf first second.]. According to this, if f is *fully amenable* at $\bar{x}$, then for all x and v such that x is sufficiently close to $\bar{x}$ and $v \in \partial f(\bar{x})$, ∂f is proto-differentiable at x for v, hence too, $(\partial f)^{-1}$ is proto-differentiable at v for x. Moreover, the proto-derivative mappings for f are then themselves the subgradient mappings for certain second-derivative function associated with f. Full amenability, a term not actually adopted until later, in [.rockaf poliqu amenab.], refers to the existence of a kind of composite representation $f(x) = g\big(F(x)\big)$ locally around $\bar{x}$ (to be reviewed in Section 3).

The class of fully amenable functions is rich for applications. It includes, for instance, any function f of the form $f_0 + \delta_C$ with δ_C the indicator of a set C specified by a finite system of $\mathcal{C}^2$ constraints such that the Mangasarian-Fromovitz constraint qualification is satisfied at $\bar{x}$, and f_0 a function of class $\mathcal{C}^2$ or the pointwise maximum of a finite collection of $\mathcal{C}^2$ functions. See [.rockaf first second.], [.rockaf genera convex.], [.rockaf poliqu calcul.], [.rockaf poliqu amenab.].

In this paper we extend such criteria for proto-differentiability beyond the class of subgradient mappings themselves to "partial" subgradient mappings. Again there is motivation from questions about parametric dependence in optimization. The basic "tilt" parameterization mentioned so far in terms of a vector v can be supplemented by a vector $p \in I\!R^m$: one can consider the minimization of $f(x,p) - \langle v, x\rangle$ in x for each $(v,p) \in I\!R^n \times I\!R^d$. The focus in that setting is on the multifunction $\Gamma : I\!R^n \times I\!R^m \rightrightarrows I\!R^n$ that associates with each pair (v,p) the set of all x satisfying the necessary condition $v \in \partial_x f(x,p)$. The proto-differentiability of Γ revolves then around that of the multifunction $\partial_x f : I\!R^n \times I\!R^d \rightrightarrows I\!R^n$ which assigns to (x,p) the subgradients of the function $f(\cdot,p)$ at x.

Contrary to what one might imagine from parallels with the classical theory of derivatives, the proto-differentiability of the subgradient mapping $\partial f : I\!R^n \times I\!R^d \rightrightarrows I\!R^n \times I\!R^d$, which is assured when $f(x,p)$ is fully amenable as a function of x and p jointly, does not guarantee for fixed p the proto-differentiability of the subgradient mapping $I\!R^n \rightrightarrows I\!R^n$ associated with the function $f(\cdot,p)$. An extra assumption (a form of constraint qualification) is needed for that, cf. [.rockaf poliqu calcul¡, Cor. 3.7¿.]. Nor does the proto-differentiability of ∂f ensure that of the partial subgradient mapping $\partial_x f$. On this topic there has been no previous work, a gap which we seek here to rectify.

Our main result has the following statement, couched in a more convenient notation where $x = (x_1, x_2)$ with $x_1 \in I\!R^{n_1}$, $x_2 \in I\!R^{n_2}$, $n_1 + n_2 = n$, and employing a parametric extension of amenability which will be clarified in Section 3. We let $\partial_1 f$ stand for the multifunction that assigns to each $x = (x_1, x_2)$ the set $\partial_1 f(x)$ of subgradients v_1 of $f(\cdot, x_2)$ at x_1. We denote by proj_1 the projection mapping from $I\!R^n = I\!R^{n_1} \times I\!R^{n_2}$ onto $I\!R^{n_1}$.

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

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

Theorem 1.1 will be proved in Section 4. The proto-derivative formula in this theorem asserts that the graph of $(\partial_1 f)'_{x,v_1}$ is the union over all vectors $v = (v_1, v_2)$ of the projection on $I\!R^{n_1} \times I\!R^{n_2} \times I\!R^{n_1}$ of the graph of $(\partial f)'_{x,v}$ as a subset of $I\!R^{n_1} \times I\!R^{n_2} \times I\!R^{n_1} \times I\!R^{n_2}$. Ordinarily, operations of projection and union would not maintain the graphical property corresponding to proto-differentiability, and this is where the challenge of our result resides. Without the parametric full amenability assumption, not only the formula but even the existence of proto-derivatives could seriously be in doubt.

Parametric full amenability is fulfilled in a wide range of applications. We illustrate this in Section 5 by applying Theorem 1.1 to the perturbation of solutions to "generalized equations." Associated with any closed set $C \subset I\!R^n$ and mapping $F : I\!R^n \to I\!R^n$ is the *variational condition*

$$F(\bar{x}) + N_C(\bar{x}) \ni 0, \quad \bar{x} \in C, \tag{1.2}$$

in which $N_C(x)$ denotes the cone of normal vectors to C at a point $x \in C$ (in the sense of limiting proximal normal vectors), but is taken to be the empty set for points $x \notin C$. When C is convex, this expresses the *variational inequality* for C and F. Sensitivity analysis has generally focused on this convex case and the study of how the set

$$\Gamma(p) = \big\{ x \,\big|\, F(x,p) + N_C(x) \ni 0 \big\} \tag{1.3}$$

depends on p under assumptions on the way $F(x,p)$ depends on p. Emphasis has been given especially to establishing continuity and differentiability properties of Γ in circumstances where it turns out to be single-valued. Robinson [.robins implici class.] provides a current overview.

A different tactic was adopted by Rockafellar [.optimi rockaf proto-d set-val.] in exploring proto-differentiability separately from single-valuedness, as motivated by the fact that, in the presence of single-valuedness and a calmness condition, proto-differentiability is equivalent to one-sided directional differentiability of the kind usually targeted in this subject. It was shown that when $F(x,p)$ is $\mathcal{C}^1$ in x and p and the set C is polyhedral, the multifunction Γ giving the solutions to $F(x,p)+N_C(x)\ni v$ in their dependence on (v,p) is proto-differentiable everywhere. This line of theory was extended recently by Levy and Rockafellar [.rockaf levy sensit genera.] in allowing C to be nonpolyhedral and even nonconvex under certain conditions, and more generally allowing $N_C(x)$ to be replaced by any set $M(x,p)$, as long as the multifunction $M:\mathbb{R}^n\times\mathbb{R}^m\rightrightarrows\mathbb{R}^n$ is proto-differentiable. The attention then is on

$$\Gamma(v,p)=\big\{x\,\big|\,F(x,p)+M(x,p)\ni v\big\}. \tag{1.4}$$

Theorem 1.1 furnishes a broad class of new examples of proto-differentiable multifunctions M which serve these purposes. We illustrate this in Section 5 with the case of $M(x,p)=N_{C(p)}$ for a variable set $C(p)$, not necessarily convex, specified by constraints in which p appears as a parameter and a standard constraint qualification is satisfied.

2. Proto-Differentiation

In this paper, we use the concept of Painlevé-Kuratowski set convergence. The *inner set limit* of a parameterized family of sets $\{G_t\}_{t>0}$ in $\mathbb{R}^N$ is the set of points η such that for *every* sequence $t_k\searrow 0$, η is a limit of a sequence of points $\eta_k\in G_{t_k}$. The *outer set limit* of the family is the set of points η such that for *some* sequence $t_k\searrow 0$, η is the limit of a sequence of points $\eta_k\in G_{t_k}$. When the inner and outer set limits are the same set G, the *limit* exists; then G_t *converges* to G as $t\searrow 0$. In our framework, this will be applied to sets that are the graphs of multifunctions.

For a multifunction $\Gamma:\mathbb{R}^m\rightrightarrows\mathbb{R}^n$ and a pair $(\bar w,\bar z)$ in $\operatorname{gph}\Gamma$, i.e., with $\bar z\in\Gamma(\bar w)$, the graph of the difference quotient mapping $(\Delta_t\Gamma)_{\bar w,\bar z}$ is $t^{-1}\big[\operatorname{gph} g-(\bar w,\bar z)\big]$. The multifunction $\Gamma'^{+}_{\bar w,\bar z}:\mathbb{R}^m\rightrightarrows\mathbb{R}^n$ having as its graph the outer limit of the sets $\operatorname{gph}(\Delta_t\Gamma)_{\bar w,\bar z}$ as $t\searrow 0$ is the *outer graphical derivative* of Γ at $\bar w$ for $\bar z$. Similarly, the *inner graphical derivative* $\Gamma'^{-}_{\bar w,\bar z}:\mathbb{R}^m\rightrightarrows\mathbb{R}^n$ corresponds to the inner limit. Proto-differentiability of Γ at $\bar w$ for $\bar z$ is defined to be the case where the outer and inner derivatives agree, the common mapping being then the *proto-derivative* $\Gamma'_{\bar w,\bar z}=\Gamma'^{+}_{\bar w,\bar z}=\Gamma'^{-}_{\bar w,\bar z}$, cf. Rockafellar [.rockaf nonsmo parame.].

When Γ is single-valued at $\bar w$, i.e., the set $\Gamma(\bar w)$ is just a singleton $\{\bar z\}$, the notation $\Gamma'_{\bar w,\bar z}$ can be simplified to $\Gamma'_{\bar w}$. When Γ is single-valued on some neighborhood of $\bar w$, it is Bdifferentiable as defined by Robinson [.robins local feasible.] if and only if it is continuous at $\bar{w}$ and proto-differentiable there with $\Gamma'_{\bar{w}}$; then one has the local expansion

$$\Gamma(w) \;=\; \Gamma(\bar{w}) + \Gamma'_{\bar{w}}(w - \bar{w}) + o\big(|w - \bar{w}|\big)$$

(cf. Levy and Rockafellar [.rockaf levy nonuniq.] for details). As seen from this angle, proto-differentiability is the natural extension to set-valued mappings of the one-sided directional differentiability property favored for ordinary single-valued mappings. If the aim in some application is to show that a solution set depending on parameters is actually a singleton whose perturbations exhibit B-differerentiability in those parameters, that can be accomplished by establishing on the one hand that the solution mapping in the set-valued sense is proto-differentiable, and on the other hand that Γ is actually single-valued around $\bar{w}$ and calm at $\bar{w}$.

An alternative description of outer derivatives, inner derivatives, and proto-derivatives will be helpful. For this we need a notion of one-sided differentiability of "arcs," a term we employ for vector-valued functions of a single variable without insisting on any preconditions, not even that continuity is necessarily present. For any arc $w : [0, \tau) \to I\!R^m$ the *right derivative of* w *at* 0 is the limit

$$w'_+(0) := \lim_{t \searrow 0} \frac{w(t) - w(0)}{t},$$

when this limit exists. Then, of course, at least $w(t) \to w(0)$ as $t \searrow 0$.

Proposition 2.1. *For a multifunction $\Gamma : I\!R^m \to I\!R^n$ and any pair $(\bar{w}, \bar{z})$ with $\bar{z} \in \Gamma(\bar{w})$, the outer graphical derivative of Γ at $\bar{w}$ for $\bar{z}$ is given by*

$$\Gamma'^{+}_{\bar{w},\bar{z}}(\omega) := \left\{ \zeta \;\middle|\; \begin{array}{l} \text{there exist sequences } \omega_k \to \omega,\ t_k \searrow 0, \text{ and} \\ \zeta_k \to \zeta \text{ such that } \bar{z} + t_k\zeta_k \in \Gamma(\bar{w} + t_k\omega_k) \end{array} \right\} \text{ for all } \omega,$$

whereas the inner graphical derivative of Γ at $\bar{w}$ for $\bar{z}$ is given by

$$\Gamma'^{-}_{\bar{w},\bar{z}}(\omega) := \left\{ \zeta \;\middle|\; \begin{array}{l} \text{there exist arcs } w : [0,\tau) \to I\!R^m \text{ and } z : [0,\tau) \to I\!R^n \\ \text{with } w(0) = \bar{w},\ z(0) = \bar{z},\ w'_+(0) = \omega, \text{ and } z'_+(0) = \zeta, \\ \text{and such that } z(t) \in \Gamma(w(t)) \text{ for all } t \in [0,\tau) \end{array} \right\} \text{ for all } \omega.$$

The inclusion $\Gamma'^{+}_{\bar{w},\bar{z}}(\omega) \supset \Gamma'^{-}_{\bar{w},\bar{z}}(\omega)$ is automatic. Proto-differentiability of Γ at $\bar{w}$ for $\bar{z}$ is the case where the opposite inclusion holds for every ω as well.

Proof. This is immediate from the definitions. □

Another fact that will be useful concerns the behavior of graphical derivatives under projection.

Proposition 2.2. *For a multifunction* $\Gamma : I\!R^m \rightrightarrows I\!R^n = I\!R^{n_1} \times I\!R^{n_2}$*, let* $\operatorname{proj}_1 \Gamma : I\!R^m \rightrightarrows I\!R^{n_1}$ *assign to each* w *the image of* $\Gamma(w)$ *under the projection* $\operatorname{proj}_1 : I\!R^n \to I\!R^{n_1}$*. Then whenever* $\bar{z}_1 \in (\operatorname{proj}_1 \Gamma)(\bar{w})$ *one has*

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

Proof. This is elementary from Proposition 2.1. □

One of the hardships we have to surmount is that the inclusions in Proposition 2.2 go in the same direction. If the inclusion for the outer graphical derivative were the opposite, we would be able to argue that equality of the inner and outer derivatives for Γ implied equality for $\operatorname{proj}_1 \Gamma$, or in other words, that proto-differentiability is preserved in passing from Γ to $\operatorname{proj}_1 \Gamma$. But that is not the reality of the subject. On the abstract level of general mappings, progress toward obtaining usable criteria for $\operatorname{proj}_1 \Gamma$ to inherit proto-differentiability from Γ is thwarted at every turn. A restriction to some special class of mappings, hopefully still ample for good applications, is essential. This is where the subgradient mappings of fully amenable functions come in.

3. Fully Amenable Functions

The class of fully amenable functions is defined in terms of special convex functions called "piecewise linear-quadratic" functions. A function $g : I\!R^m \to \overline{I\!R} := [-\infty, \infty]$ is *piecewise linear-quadratic (p.l.q.)* if its effective domain, $\operatorname{dom} g$, is the union of a finite number of (convex) polyhedral sets over which the function is linear or quadratic (see [.rockaf first nonlinear progra.]). Examples of convex p.l.q. functions include affine functions, the pointwise maximum of finitely many affine functions, the indicators of (convex) polyhedral sets, positive-semidefinite quadratic functions, and functions of the form dist^2_C for a polyhedral set C, where dist_C is the function giving for each point its distance from C. Any nonnegative linear combination of convex p.l.q. functions is again convex p.l.q., and so too is the conjugate of any such function under the Legendre-Fenchel transform. An example of the latter is a penalty function of the type

$$\rho_{Y,Q}(u) = \sup_{y \in Y} \big\{ \langle y, u \rangle - \tfrac{1}{2} \langle y, Qy \rangle \big\} \tag{3.1}$$

for a polyhedral set Y and positive-*semi*definite, symmetric matrix Q. This is dual to $\delta_Y(y) + \frac{1}{2}\langle y, Qy \rangle$. Rockafellar [.rockaf linear-quadratic program.] has shown that such functions can be used to represent many expressions typically seen in optimization problems. The “penalty” label alludes to the fact that quite general penalties for violating constraints (or rewards for staying within them) can be formulated as functions of type (3.1).

A function $f : I\!R^n \to \overline{I\!R}$ is *fully amenable at* $\bar{x}$, a point where $f(\bar{x})$ is finite, if on some neighborhood of $\bar{x}$ it has a representation $f(x) = g\big(F(x)\big)$ in which F is a $\mathcal{C}^2$ mapping from this neighborhood into a space $I\!R^m$, while g is a convex p.l.q. function on $I\!R^m$ such that the following constraint qualification is fulfilled at $\bar{x}$:

$$0 \in \operatorname{int}\Big(\operatorname{dom} g - \big[F(\bar{x}) + \nabla F(\bar{x}) I\!R^n\big]\Big), \tag{3.2}$$

or equivalently in terms of normal vectors to $\operatorname{dom} g$,

$$\left.\begin{array}{r} y \in N_{\operatorname{dom} g}\big(F(\bar{x})\big) \\ \nabla F(\bar{x})^* y = 0 \end{array}\right\} \quad \Longrightarrow \quad y = 0. \tag{3.3}$$

Here $\nabla F(\bar{x})$ denotes the Jacobian matrix for F at $\bar{x}$, while $\nabla F(\bar{x})^*$ is its transpose. The set $F(\bar{x}) + \nabla F(\bar{x}) I\!R^n$ in (3.2) is the range of the “linearization” of F at $\bar{x}$. The interiority condition means that the convex set $\operatorname{dom} g$ cannot be separated from this affine range set, which in terms of normal vectors translates to (3.3).

A function that is fully amenable at a point $\bar{x}$ is in fact fully amenable at every point in a neighborhood of $\bar{x}$. A function that is fully amenable at all points in its effective domain is simply called a *fully amenable function.* (By itself, *amenability* merely assumes in the setting of Definition 3.1 that F is $\mathcal{C}^1$, whereas g is lsc, proper and convex; *strong* amenability has such a general g but a $\mathcal{C}^2$ mapping F.)

Some important examples of fully amenable functions have already been mentioned in the introduction. All convex p.l.q. functions and all $\mathcal{C}^2$ functions fit the definition too as special cases. A set C is fully amenable if its indicator δ_C is fully amenable. Rules for preserving full amenability under operations like addition and composition are also known, along with formulas for the proto-derivatives for subgradient mappings constructed in various circumstances. For more on the topic, see [.rockaf first second.], [.rockaf genera convex.], [.rockaf poliqu calcul.], [.rockaf poliqu amenab.], [.rockaf poliqu variat.].

Definition 3.1. *Consider a function* $f : I\!R^{n_1} \times I\!R^{n_2} \to \overline{I\!R}$ *and a point* $\bar{x} = (\bar{x}_1, \bar{x}_2)$ *with* $f(\bar{x})$ *finite. We shall say that* $f(x_1, \bar{x}_2)$ *is fully amenable in* x_1 *at* $\bar{x}_1$ *with compatible parameterization in* x_2 *at* $\bar{x}_2$ *if on some neighborhood of* $\bar{x}$ *there is a representation* $f(x) =$

$f(x_1,x_2) = g\big(F(x_1,x_2)\big)$ in which F is a $\mathcal{C}^2$ mapping from this neighborhood into a space $I\!R^m$, while g is a convex p.l.q. function on $I\!R^m$, such that the following constraint qualification is fulfilled:

$$0 \in \operatorname{int}\Big(\operatorname{dom} g - \big[F(\bar{x}) + \nabla_1 F(\bar{x}) I\!R^{n_1}\big]\Big), \tag{3.4}$$

or equivalently in terms of normal vectors to $\operatorname{dom} g$,

$$\left.\begin{array}{r} y \in N_{\operatorname{dom} g}\big(F(\bar{x})\big) \\ \nabla_1 F(\bar{x})^* y = 0 \end{array}\right\} \quad \Longrightarrow \quad y = 0. \tag{3.5}$$

Here $\nabla_1 F(\bar{x}) = \nabla_1 F(\bar{x}_1,\bar{x}_2)$ is the Jacobian matrix for $F(x_1,\bar{x}_2)$ with respect to x_1 at $\bar{x}_1$, while $\nabla_1 F(\bar{x})^*$ is its transpose.

The equivalence claimed between (3.4) and (3.5) is clear from the interpretation of (3.4) as characterizing the impossibility of separating the convex set $\operatorname{dom} g$ from the affine set $F(\bar{x}) + \nabla_1 F(\bar{x}) I\!R^{n_1}$ even improperly.

The distinction between this concept and the full amenability of f at $\bar{x}$ is that the constraint qualification involves not the range of the entire linearization of F at $\bar{x}$, but only its linearization in the x_1 component with the x_2 component fixed at $\bar{x}_2$. This in general is a smaller range set, so the condition is a stronger one. It is also stronger than merely requiring the function $f(\cdot,\bar{x}_2)$ to be fully amenable at $\bar{x}_1$. That would involve essentially the same constraint qualification as (3.4) but without any provision for the role of x_2; it would mean the existence, perhaps only for the single value $\bar{x}_2$, of a representation $f(x_1,\bar{x}_2) = g\big(F_1(x_1)\big)$ in which F_1 is a $\mathcal{C}^2$ mapping from $I\!R^{n_1}$ to $I\!R^m$ and g is a convex p.l.q. function on $I\!R^m$ such that

$$0 \in \operatorname{int}\Big(\operatorname{dom} g - \big[F_1(\bar{x}_1) + \nabla F_1(\bar{x}_1) I\!R^{n_1}\big]\Big). \tag{3.6}$$

The condition in Definition 3.1 guarantees that this is satisfied by the choice $F_1(x_1) = F(x_1,\bar{x}_2)$. But the argument cannot be made in the opposite direction; there is no automatic way to go from a representation $f(x_1,\bar{x}_2) = g\big(F_1(x_1)\big)$ to one of the form $f(x_1,x_2) = g\big(F(x_1,x_2)\big)$ involving values x_2 around $\bar{x}_2$ and $F(x_1,x_2)$ not only $\mathcal{C}^2$ in x_1 but in (x_1,x_2). It is this parametric feature that Definition 3.1 requires.

These observations are summarized as follows, along with the subgradient formulas that will be basic to our task.

Proposition 3.2. *Consider a function* $f : I\!R^{n_1} \times I\!R^{n_2} \to \overline{I\!R}$ *and a point* $\bar{x} = (\bar{x}_1, \bar{x}_2)$*. If* $f(x_1, \bar{x}_2)$ *is fully amenable in* x_1 *at* $\bar{x}_1$ *with compatible parameterization in* x_2 *at* $\bar{x}_2$*, then in particular*

(a) f *is fully amenable at* $\bar{x}$*,*

(b) $f(\cdot, \bar{x}_2)$ *is fully amenable at* $\bar{x}_1$*.*

Moreover, with respect to any local representation of f *as in Definition 3.1, one has*

$$\begin{aligned} \partial f(\bar{x}) &= \nabla F(\bar{x})^* \partial g\big(F(\bar{x})\big) = \big\{(\nabla_1 F(\bar{x})^* y, \nabla_2 F(\bar{x})^* y) \,\big|\, y \in \partial g\big(F(\bar{x})\big)\big\}, \\ \partial_1 f(\bar{x}) &= \nabla_1 F(\bar{x})^* \partial g\big(F(\bar{x})\big) = \big\{\nabla_1 F(\bar{x})^* y \,\big|\, y \in \partial g\big(F(\bar{x})\big)\big\}. \end{aligned}$$

Proof. The truth of (a) and (b) has just been observed. In Rockafellar [.rockaf first second.] it was proved that when a function f is represented as $g \circ F$ in the manner dictated by "amenability" (whether full or not), its subgradient mapping satisfies the identity $\partial f(x) = \nabla F(x)^* \partial g(F(x))$. The two formulas follow from this, the first for the whole of f as a function of (x_1, x_2) and the second by application to $f_1(x_1) = g\big(F_1(x_1)\big)$ with $F_1(x_1) = F(x_1, \bar{x}_2)$. □

We note further that parametric full amenability, like all other versions of amenability, is a property that, if it holds at a certain point, holds locally around that point.

Proposition 3.3. *When the parametric full amenability condition in Definition 3.1 is satisfied at* $\bar{x}$*, it is satisfied also by all* x *in some neighborhood of* $\bar{x}$ *relative to* $\operatorname{dom} f$*.*

Proof. This is clear from form (3.4) of the constraint qualification in view of the continuity of F and $\nabla_1 F$ and the fact that, for any convex set C, if $y_k \in N_C(u_k)$ with $u_k \to u \in C$ and $y_k \to y$, then $y \in N_C(u)$. □

4. Subgradient Projection

Parametric full amenability will support the development, for partial subgradient mappings, of an inclusion for projected outer derivatives that complements the elementary one in Proposition 2.2 by going in the opposite direction. This will be the foundation on which we shall be able to establish our main result, Theorem 1.1.

In [.sun.], Sun showed that the subgradient mapping associated with any proper, convex, p.l.q. function is *Robinson polyhedral* [.robins continu multif.]: its graph is the union of finitely many polyhedral sets. This will be crucial. The technical property it makes available is described next.

Lemma 4.1. *Consider $f : I\!R^n \to \overline{I\!R}$ with $I\!R^n = I\!R^{n_1} \times I\!R^{n_2}$ and a point $\bar{x} = (\bar{x}_1, \bar{x}_2)$ with $f(\bar{x})$ finite. Suppose that $f(x_1, \bar{x}_2)$ is fully amenable in x_1 at $\bar{x}_1$ with compatible parameterization in x_2 at $\bar{x}_2$, and let $f = g \circ F$ be a representation around $\bar{x}$ as demanded by this property, with $F : I\!R^n \to I\!R^m$ and $g : I\!R^m \to \overline{I\!R}$.*

Let the subspace $M \subset I\!R^m \times I\!R^m$ be defined by $M = I\!R^m \times \nabla_1 F(\bar{x}) I\!R^{n_1}$. Then for any point $\bar{y} \in \partial g(F(\bar{x}))$ there exist balls B_1 and B_2 centered at the origin of $I\!R^m \times I\!R^m$ such that the set $G := \operatorname{gph} \partial g - (F(\bar{x}), \bar{y})$ and its projection $\operatorname{proj}_M G$ on M satisfy

$$tB_1 \cap \operatorname{proj}_M G \subset \operatorname{proj}_M \big\{ G \cap tB_2 \big\} \text{ for all } t > 0 \text{ sufficiently small.} \tag{4.1}$$

Proof. As defined, we have $G = \big\{ (u, y) \in I\!R^m \times I\!R^m \,\big|\, y \in \partial g\big(F(\bar{x}) + u\big) - \bar{y} \big\}$, whereas $\operatorname{proj}_M G$ consists of the pairs $(u, z) \in I\!R^m \times I\!R^m$ such that

$$\begin{cases} \text{there exist } y \in \partial g\big(F(\bar{x}) + u\big) - \bar{y} \text{ and } \xi_1 \in I\!R^{n_1} \\ \text{with } z = \nabla_1 F(\bar{x}) \xi_1 \text{ and } y - z \perp \nabla_1 F(\bar{x}) I\!R^{n_1}, \end{cases} \tag{4.2}$$

where obviously

$$y - z \perp \nabla_1 F(\bar{x}) I\!R^{n_1} \iff \langle \nabla_1 F(\bar{x})^* y, \xi_1 \rangle = \langle \nabla_1 F(\bar{x})^* z, \xi_1 \rangle \text{ for all } \xi_1 \in I\!R^{n_1}. \tag{4.3}$$

Because the set G contains the origin and is the union of finitely many polyhedral sets, we are assured that $\operatorname{proj}_M G$ likewise contains the origin and is the union of finitely many polyhedral sets. Hence there exists a ball B_1 about the origin of $I\!R^m \times I\!R^m$ such that

$$t\big(\operatorname{proj}_M G \cap B_1 \big) = \operatorname{proj}_M \big(G \cap tB_1 \big) \text{ for all } t > 0 \text{ sufficiently small.} \tag{4.4}$$

The set $\operatorname{dom} g - \big[F(\bar{x}) + \nabla_1 F(\bar{x}) I\!R^{n_1} \big]$, which the constraint qualification (3.4) in the definition of parametric full amenability requires to contain the origin in its interior, can be expressed in terms of the level sets $\operatorname{lev}_\alpha g = \big\{ u \,\big|\, g(u) \le \alpha \big\}$ as the union of the convex sets $\operatorname{lev}_\alpha g - \big[F(\bar{x}) + \nabla_1 F(\bar{x}) X_1 \big]$ over all $\alpha \in I\!R$ and closed balls X_1 centered at the origin of $I\!R^{n_1}$. In fact it can be expressed as the increasing union of countably many of such sets, which are closed. By the Baire category theorem, such a union cannot have the origin in its interior unless one of the sets comprising it already has the origin in its interior. Thus, for some $\alpha \in I\!R$ and closed ball X_1 at the origin of $I\!R^{n_1}$, there is a closed ball U at the origin of $I\!R^m$ such that

$$2U \subset \Big(\operatorname{lev}_\alpha g - \big[F(\bar{x}) + \nabla_1 F(\bar{x}) X_1 \big] \Big). \tag{4.5}$$

Let β be the minimum of $g\big(F(\bar{x})+u\big)$ over all $u\in U$; this is finite because g is lsc and proper with $g\big(F(\bar{x})\big)$ finite. By shrinking the ball B_1 in (4.4) somewhat if necessary, we can suppose that pairs $(u,z)\in \mathrm{proj}_M\, G\cap B_1$ all have $u\in U$. Then for any such pair (u,z) and corresponding y as in (4.2), and for arbitrary $u'\in U$, there exists through (4.5) some $\xi_1\in X_1$ such that $u+u'+F(\bar{x})+\nabla_1 F(\bar{x})\xi_1\in \mathrm{lev}_\alpha\, g$, i.e., $g\big(F(\bar{x})+u+u'+\nabla_1 F(\bar{x})\xi_1\big)\le\alpha$. But because g is convex and $\bar{y}+y\in\partial g\big(F(\bar{x})+u\big)$ we have the subgradient inequality

$$g\big(F(\bar{x})+u+u'+\nabla_1 F(\bar{x})\xi_1\big)\ \ge\ g\big(F(\bar{x})+u\big)+\big\langle \bar{y}+y, u'+\nabla_1 F(\bar{x})\xi_1\big\rangle,$$

where $g\big(F(\bar{x})+u\big)\ge\beta$ by the selection of β, and $\langle\nabla_1 F(\bar{x})^* y,\xi_1\rangle=\langle\nabla_1 F(\bar{x})^* z,\xi_1\rangle$ by (4.3). This tells us that

$$\big\langle \bar{y}+y,u'\big\rangle\ \le\ \alpha-\beta-\big\langle\bar{y}+z,\nabla_1 F(\bar{x})\xi_1\big\rangle.$$

The right side of this inequality has an upper bound relative to all possible instances of $(u,z)\in\mathrm{proj}_M\, G\cap\mathcal{B}_1$ and $\xi_1\in X_1$, so the fact that u' on the left side ranges over the ball U implies the existence of an upper bound on the norm of $\bar{y}+y$ that does not depend on the particular (u,z) under investigation. Thus, there is a ball B_2 around the origin in $\mathbb{R}^m\times\mathbb{R}^m$ such that whenever $(u,z)\in\mathrm{proj}_M\, G\cap\mathcal{B}_1$ and y is a corresponding vector as in (4.2), one has $(u,y)\in B_2$. In other words,

$$\mathrm{proj}_M\, G\cap B_1\subset \mathrm{proj}_M\,\big(G\cap B_2\big). \tag{4.6}$$

But again, since G contains the origin and is the union of finitely many polyhedral sets, we have in parallel with (4.4) that

$$t\big(G\cap B_2\big)=G\cap tB_2\ \text{ for all }\ t>0\ \text{ sufficiently small.}$$

In combination with (4.6) and (4.4), this produces the desired estimate (4.1). □

Proposition 4.2. *Consider $f:\mathbb{R}^n\to\overline{\mathbb{R}}$ with $\mathbb{R}^n=\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ and a point $\bar{x}=(\bar{x}_1,\bar{x}_2)$ with $f(\bar{x})$ finite. Suppose that $f(x_1,\bar{x}_2)$ is fully amenable in x_1 at $\bar{x}_1$ with compatible parameterization in x_2 at $\bar{x}_2$. Let $\bar{v}_1\in\partial_1 f(\bar{x})$. Then*

$$(\partial_1 f)'^{+}_{\bar{x},\bar{v}_1}(\xi)\ \subset\ \bigcup\nolimits_{\bar{v}}\Big\{\,\mathrm{proj}_1(\partial f)'^{+}_{\bar{x},\bar{v}}(\xi)\,\Big|\,\mathrm{proj}_1\,\bar{v}=\bar{v}_1\Big\}\ \textit{ for all }\xi.$$

Proof. Fix any $\theta_1\in(\partial_1 f)'^{+}_{\bar{x},\bar{v}_1}(\xi)$. Our task is to prove the existence of $\theta_2\in\mathbb{R}^{n_2}$ and $\bar{v}_2\in\mathbb{R}^{n_2}$ such that the vector $\bar{v}=(\bar{v}_1,\bar{v}_2)$ belongs to $\partial f(\bar{x})$ and the vector $\theta=(\theta_1,\theta_2)$ belongs to $(\partial f)'^{+}_{\bar{x},\bar{v}}(\xi)$.

From the definitions we know there are sequences $t_k \searrow 0$, $\xi_k \to \xi$ and $\theta_{1k} \to \theta_1$ with

$$\bar{v}_1 + t_k\theta_{1k} \in \partial_1 f(\bar{x} + t_k\xi_k). \tag{4.7}$$

Let $f = g \circ F$ be a representation around $\bar{x}$ as specified in Definition 3.1, with $F : I\!R^n \to I\!R^m$ and $g : I\!R^m \to \overline{I\!R}$. The facts in Propositions 3.2 and 3.3 are then available. In particular, we have in some neighborhood X of $\bar{x}$ that $\partial_1 f(x) = \nabla_1 F(x)^* \partial g\big(F(x)\big)$ for all $x \in X$. On this basis we can write (4.7) as

$$\bar{v}_1 + t_k\theta_{1k} = \nabla_1 F(\bar{x} + t_k\xi_k)^* y_k \;\text{ with }\; y_k \in \partial g\big(F(\bar{x} + t_k\xi_k)\big). \tag{4.8}$$

The sequence $\{y_k\}$ introduced in this manner must be bounded, for if not we could reduce to the case where $|y_k| \to \infty$, and where the vectors $\hat{y}_k = y_k/|y_k|$ converge to some $y \neq 0$. Then $\nabla_1 F(\bar{x})^* y = 0$ from (4.8), whereas $y \in N_{\operatorname{dom} g}\big(F(\bar{x})\big)$ out of the fact that $\hat{y}_k \in \partial(\varepsilon_k g)\big(F(\bar{x}+t_k\xi_k)\big)$ for $\varepsilon_k = 1/|y_k|$. This would contradict (3.4), the alternate form of the constraint qualification underlying parametric full amenability. Hence we can suppose $\{y_k\}$ converges to some $\bar{y}$, necessarily with

$$\bar{y} \in \partial g\big(F(\bar{x})\big), \qquad \nabla F_1(\bar{x})^* \bar{y} = \bar{v}_1. \tag{4.9}$$

At the same time we can translate the subgradient condition in (4.8) into

$$\bar{y} + t_k\eta_k \in \partial g\big(F(\bar{x}) + t_k\omega_k\big) \;\text{ for }\; \eta_k := \frac{y_k - \bar{y}}{t_k}, \tag{4.10}$$

where

$$\omega_k := \frac{F(\bar{x} + t_k\xi_k) - F(\bar{x})}{t_k} \;\to\; \omega := \nabla F(\bar{x})\xi. \tag{4.11}$$

Then we can write

$$\theta_{1k} = \nabla_1 F(\bar{x})^* \eta_k + d_{1k} \;\text{ for }\; d_{1k} := \left[\frac{\nabla_1 F(\bar{x} + t_k\xi_k)^* - \nabla_1 F(\bar{x})^*}{t_k}\right] y_k \tag{4.12}$$

with the knowledge that, because F is a $\mathcal{C}^2$ mapping, the sequence $\{d_{1k}\}$ converges to a certain vector d_1.

The rest of the proof hinges on the possible unboundedness of the sequence $\{\eta_k\}$ and how to get around that, if need be, by a shift in the sequence $\{y_k\}$ selected so far. We shall rely on the decomposition

$$I\!R^m = N + N^\perp \;\text{ for }\; N := \nabla_1 F(\bar{x}) I\!R^{n_1}$$

and the mapping proj_N that gives the orthogonal projection onto the subspace N. Note that two vectors η and η' have $\operatorname{proj}_N \eta = \operatorname{proj}_N \eta'$ if and only if $\nabla_1 F(\bar{x})^* \eta = \nabla_1 F(\bar{x})^* \eta'$.

For each k let $\zeta_k = \operatorname{proj}_N \eta_k$. Then $\zeta_k = \operatorname{proj}_N \zeta_k$ as well, so that

$$\nabla_1 F(\bar{x})^* \eta_k = \nabla_1 F(\bar{x})^* \zeta_k \ \text{ for all } k. \tag{4.13}$$

It will be shown that the sequence $\{\zeta_k\} \subset N$ is bounded. For this it suffices to demonstrate that for every $w = \nabla_1 F(\bar{x}) x_1$ in N the sequence $\{\langle \zeta_k, w \rangle\}$ is bounded. But from (4.12) and (4.13) we have $\big\langle \zeta_k, \nabla_1 F(\bar{x}) x_1 \big\rangle = \big\langle \theta_{1k} - d_{1k}, x_1 \big\rangle \to \big\langle \theta_1 - d_1, x_1 \big\rangle$, so this is assured.

As in Lemma 4.1, let $G = \operatorname{gph} g - \big(F(\bar{x}), \bar{y}\big)$ and consider the subspace $M = I\!R^m \times N$ of $I\!R^m \times I\!R^m$. By (4.10) we have $t_k(\omega_k, \eta_k) \in G$, whereas $t_k(\omega_k, \zeta_k) \in \operatorname{proj}_M G$. Let B_1 and B_2 be balls with the property provided by Lemma 4.1. Because the sequence $\{(\omega_k, \eta_k)\}$ is bounded, it lies in λB_1 for some $\lambda > 0$. For k sufficiently large, therefore, we have

$$t_k(\omega_k, \zeta_k) \ \in \ \operatorname{proj}_M G \cap t_k \lambda B_1 \ \subset \ \operatorname{proj}_M (G \cap t_k \lambda_k B_2),$$

which means that $\zeta_k = \operatorname{proj}_N \eta'_k$ for certain vectors η'_k such that $t_k(\omega_k, \eta'_k) \in G$ and $(\omega_k, \eta'_k) \in t_k \lambda B_2$. Then

$$\bar{y} + t_k \eta'_k \in \partial g\big(F(\bar{x}) + t_k \omega_k\big) = \partial g\big(F(\bar{x} + t_k \xi_k)\big)$$

and the sequence $\{\eta'_k\}$ is bounded. Passing to subsequences if necessary, we can suppose henceforth that η'_k converges to a certain η'.

The vectors $y'_k := \bar{y} + t_k \eta'_k$, like $y_k = \bar{y} + t_k \eta_k$, converge to $\bar{y}$. Because $\operatorname{proj}_N \eta_k = \operatorname{proj}_N \eta'_k$, we have $\nabla_1 F(\bar{x})^* \eta_k = \nabla_1 F(\bar{x})^* \eta'_k$ and consequently $\nabla_1 F(\bar{x})^* y_k = \nabla_1 F(\bar{x})^* y'_k$. We can therefore reconstitute (4.8) as

$$\bar{v}_1 + t_k \theta'_{1k} = \nabla_1 F(\bar{x} + t_k \xi_k)^* y'_k \ \text{ with } y'_k \in \partial g\big(F(\bar{x} + t_k \xi_k)\big) \tag{4.14}$$

by observing that

$$\nabla_1 F(\bar{x} + t_k \xi_k)^* y'_k - \nabla_1 F(\bar{x} + t_k \xi_k)^* y_k = t_k \big[\nabla_1 F(\bar{x} + t_k \xi_k)^* - \nabla_1 F(\bar{x})^*\big](y'_k - y_k)$$

and defining

$$\theta'_{1k} := \theta_{1k} - \left[\frac{\nabla_1 F(\bar{x} + t_k \xi_k)^* - \nabla_1 F(\bar{x})^*}{t_k}\right](y'_k - y_k).$$

Since F is $\mathcal{C}^2$ and $y'_k - y_k \to 0$, the sequence $\{\theta'_{1k}\}$, like $\{\theta_{1k}\}$, converges to θ_1.

If we let $\bar{v}_2 = \nabla_2 F(\bar{x})^* \bar{y}$, setting $\bar{v} = (\bar{v}_1, \bar{v}_2)$, and define the sequence $\{\theta'_{2k}\}$ by

$$\theta'_{2k} = \frac{\nabla_2 F(\bar{x} + t_k \xi_k)^* y'_k - \bar{v}_2}{t_k}$$

then we obtain from Proposition 3.2 that the pairs $\theta'_k := (\theta'_{1k}, \theta'_{2k})$ satisfy

$$\bar{v} + t_k \theta'_k \in \partial f(\bar{x} + t_k \xi_k) \tag{4.15}$$

for all k. The definition of θ'_{2k} can be rewritten as

$$\theta'_{2k} = \nabla_2 F(\bar{x})^* \eta'_k + \left[\frac{\nabla_2 F(\bar{x} + t_k \xi_k)^* - \nabla_2 F(\bar{x})^*}{t_k} \right] \bar{y},$$

and this expression converges to a certain θ_2 because $\eta'_k \to \eta$, $y'_k \to \bar{y}$, and F is $\mathcal{C}^2$. Then θ'_k converges to $\theta := (\theta_1, \theta_2)$, and by (4.15) we have $\theta \in (\partial f)'^{+}_{\bar{x},\bar{v}}(\xi)$, as required. □

Proof of Theorem 1.1. The first of the formulas in this theorem was already covered by Proposition 3.2. The second now follows from combining the general inner derivative inclusion of Proposition 2.2, as specialized to $\Gamma = \partial_1 f$, with the outer derivative inclusion of Proposition 4.2. □

5. Application to Parameterized Variational Conditions

To illustrate some of the content of our results, we turn now to the sensitivity analysis of variational conditions in which the underlying set can depend on the parameters. For this purpose we take the formulation

$$C(p) = \{ x \mid G(x,p) \in D \} \subset I\!R^n \tag{5.1}$$

for a set $D \subset I\!R^m$ and a mapping $G : I\!R^n \times I\!R^d \to I\!R^m$. A standard example is

$$D = \{ u = (u_1, \ldots, u_m) \mid u_1 \le 0, \ldots, u_s \le 0, u_{s+1} = 0, \ldots, u_m = 0 \}, \tag{5.2}$$

in which case, with $G(x,p) = \big(g_1(x,p), \ldots, g_m(x,p)\big)$, we would have $C(p)$ consisting of all $x \in I\!R^n$ that satisfy

$$g_i(x,p) \begin{cases} \le 0 & \text{for } i = 1, \ldots, s, \\ = 0 & \text{for } i = s+1, \ldots, m. \end{cases} \tag{5.3}$$

More generally, for instance, D could be a box giving upper and lower bounds on the components of x (with infinities as a special case). Note that in these examples D happens to be a (convex) *polyhedral* set.

It will be useful to think of $C(p)$ for each p as the x-section at p of the set

$$E = \big\{(x,p) \in I\!R^n \times I\!R^d \,\big|\, G(x,p) \in D\big\}. \tag{5.4}$$

Theorem 5.1. *For a variable set $C(p)$ as in (5.1) with D polyhedral and the mapping G of class $\mathcal{C}^2$, consider a parameter vector $\bar{p}$ and a point $\bar{x} \in C(\bar{p})$. Suppose the following constraint qualification is fulfilled at $\bar{x}$:*

$$\left.\begin{array}{r} y \in N_D\big(G(\bar{x},\bar{p})\big) \\ \nabla_x G(\bar{x},\bar{p})^* y = 0 \end{array}\right\} \quad\Longrightarrow\quad y = 0. \tag{5.5}$$

Then for x and p sufficiently near to $\bar{x}$ and $\bar{p}$ with $x \in C(p)$ one has

$$N_{C(p)}(x) = \big\{v = \nabla_x G(x,p)^* y \,\big|\, y \in N_D\big(G(x,p)\big)\big\}, \tag{5.6}$$

and the multifunction $M : (x,p) \mapsto N_{C(p)}(x)$ is proto-differentiable at (x,p) for every vector $v \in N_{C(p)}(x)$.

Proof. Let $f = \delta_E$ for the set E in (5.4). We have $f = \delta_D \circ G$, where the indicator function δ_D is convex and p.l.q. because D is polyhedral. The constraint qualification (5.5) is the one demanded by Definition 3.1 in version (3.4) in order to know that $f(x,\bar{p})$ is fully amenable in x at $\bar{x}$ with compatible parameterization in p at $\bar{p}$. Therefore, f has this property. We have $\partial_x f(x,p) = N_{C(p)}(x)$, so $M = \partial_x f$ and the desired conclusion is obtained at once from Theorem 1.1 using the normal cone formula for E in Rockafellar [.rockaf first second¡, Thm. 4.5¿.] (as applied to $f = \delta_E$). ◻

In the standard case of (5.2), the constraint qualification (5.5) is the Mangasarian-Fromovitz condition, while $N_{C(p)}(x)$ is the usual cone generated by the gradients of the active constraints in (5.3) from combinations with arbitrary coefficients relative to equality constraints but nonnegative coefficients relative to inequality constraints.

Theorem 5.2. *Consider a variable set $C(p)$ under the assumptions in Theorem 5.1 along with a $\mathcal{C}^1$ mapping $F : I\!R^n \times I\!R^d \to I\!R^n$. Define the multifunction $\Gamma : I\!R^n \times I\!R^d \to I\!R^n$ by*

$$\Gamma(v,p) := \Big\{x \,\Big|\, F(x,p) + M(x,p) \ni v\Big\}, \text{ where } M(x,p) = N_{C(p)}(x). \tag{5.7}$$

Then whenever $\bar{x} \in \Gamma(\bar{v}, \bar{p})$ and the constraint qualification (5.5) is fulfilled, one has for all (v, p) sufficiently close to $(\bar{v}, \bar{p})$ and all $x \in \Gamma(v, p)$ that Γ is proto-differentiable at (v, p) for x. Moreover

$$\begin{aligned} \Gamma'_{v,p,x}(v', p') = \Big\{ x' \,\Big|\, F'_{x,p}(x', p') + M'_{x,p,w}(x', p') \ni v' \Big\}, \ \text{where} \\ w = v - F(x,p) \ \text{and} \ F'_{x,p}(x', p') = \nabla_x F(x,p)x' + \nabla_p F(x,p)p', \end{aligned} \tag{5.8}$$

in which, for the set E in (5.4), one has

$$M'_{x,p,w}(x', p') = \bigcup_q \Big\{ v' \,\Big|\, \exists q' \ \text{with} \ (v', q') \in (N_E)'_{(x,p),(w,q)}(x', p') \Big\}. \tag{5.9}$$

Proof. We simply invoke Theorem 4.1 of Levy and Rockafellar [.rockaf levy general.] in the light of the proto-differentiability of M in the present Theorem 5.1. Formula (5.9) specializes the projection formula for proto-derivatives in Theorem 1.1 to this case. □

For the details on proto-derivatives of normal cone mappings, so as to fill in specifics for a term like $(N_E)'_{(x,p)(w,q)}(x', p')$ in (5.9), we refer to Rockafellar and Poliquin [.rockaf poliqu formula.].

Theorem 5.2 is noteworthy because it allows the parameter vector p to affect the set C along with F. To obtain proto-derivatives in such a setting without the aid of this result, the only recourse would be to work instead with E or some polyhedral set derived from D. In the case of E, as suggested by the proof of Theorem 5.1, the solution multifunction would be given by

$$\Gamma_1(v, p) = \Big\{ (x, q) \,\Big|\, \big(F(x,p), 0\big) + N_E(x,p) \ni (v, q) \Big\}.$$

This, however, would require coping with a variational condition that is not a variational inequality unless E is convex, and which therefore would anyway necessitate appealing to the extended theory in Levy and Rockafellar [.rockaf levy general.]. Also, by incorporating an additional element q as part of the solution it would deflect the sensitivity analysis into a somewhat different environment.

To revert to the standard format of the subject, one would have to pass to a variational inequality over some convex set related to D. In the case where D is a cone, as in (5.2), one could work for instance with the multifunction

$$\begin{aligned} \Gamma_2(u, v, p) = \Big\{ (x, y) \,\Big|\, F_0(x, y, p) + N_{C_0}(x, y) \ni (v, u) \Big\} \ \text{with} \\ C_0 = \mathbb{R}^n \times D^*, \quad F_0(x, y, p) = \big(F(x,p) + \nabla_x G(x,p)^* y, -G(x,p)\big), \end{aligned}$$

where D^* is the polyhedral cone polar to D (so that the condition $y \in N_D\big(G(x,p)\big)$ is equivalent to $-G(x,p) + N_{D^*}(y) \ni u$ for $u = 0$). But this would deflect the analysis still farther from its basic goal, in effect putting the emphasis on circumstances in which not just x but the multiplier vector y would be well behaved with respect to (v,p)—and also with respect to perturbations of the vector u away from $u = 0$. Previous results concerning differentiability properties of solution multifunctions have resorted to just such a reformulation, starting from a standard nonlinear programming problem where the set $C(p)$ is given by (5.3) (see [.qiu magnanti analysis.], [.kyparisis framework.], [.ralph dempe.], and [.fiacco kyparisis.] for a survey).

We remark further that the results of Levy and Rockafellar [.rockaf levy general.] also make provision for situations where $F(x,p)$ is not necessarily defined for all $p \in I\!R^m$ but only perhaps with respect to a parameter set $P \subset I\!R^m$ with boundary, for instance a box or simplex.

For other applications of the results in this paper to sensitivity analysis in optimization, see Levy and Rockafellar [.rockaf levy nonuniq.].

References

.[]