SENSITIVITY ANALYSIS OF SOLUTIONS TO GENERALIZED EQUATIONS

A. B. LEVY AND R. T. ROCKAFELLAR

ABSTRACT. Generalized equations are common in the study of optimization through nonsmooth analysis. For instance, variational inequalities can be written as generalized equations involving normal cone mappings, and have been used to represent first-order optimality conditions associated with optimization problems. Therefore, the stability of the solutions to first-order optimality conditions can be determined from the differential properties of the solutions of parameterized generalized equations. In finite-dimensions, solutions to parameterized variational inequalities are known to exhibit a type of generalized differentiability appropriate for multifunctions. Here it is shown, in a Banach space setting, that solutions to a much broader class of parameterized generalized equations are "differentiable" in a similar sense.

1. INTRODUCTION

One of the most important areas of nonsmooth analysis is the study of the sensitivity to perturbation of the solutions to parameterized optimization problems. Since it is difficult, in general, to identify actual *solutions* to optimization problems, sensitivity analysis is often carried out instead with respect to the sets of *quasi-solutions* (points satisfying a first-order necessary condition for optimality). In the case of *convex* problems, where first-order necessary conditions for optimality are sufficient as well, these quasi-solutions are *true* solutions. Since many first-order conditions for optimality can be expressed in the form of generalized equations, there is much interest in the stability of the solutions to parameterized generalized equations (see [1–5] for instance).

Robinson [3] gave an excellent survey of work, through the late 1980's, on the sensitivity analysis of generalized equations. These results can be interpreted in terms of the sensitivity to perturbations, in p, of the multifunction (set-valued mapping) G given by:

$$(1.1.1) \qquad G(p) = \{x \in C \mid -F(p, x) \in M(x)\}$$

where the sensitivities of the (single-valued) mapping F and the multifunction M are given. Robinson [3] showed that G is singled-valued and Lipschitz when F is continuous in both variables, Lipschitz in the first and "strongly Bouligand"

differentiable in the second, and M is the normal cone mapping associated with C, a closed, convex set in a Hilbert space. Under strong additional assumptions, Robinson [3] actually obtained the "Bouligand" differentiability of the necessarily single-valued G. Dontchev and Hager [4] extended Robinson's results and obtained generalized Lipschitz results for the mapping G, with M more general than a normal cone.

Further work has been done to obtain not only "continuity" results but "*differentiability*" results for the mapping G. When M is the normal cone mapping associated with the polyhedral, convex set $C \subseteq \mathbb{R}^n$, Robinson [1] proved that under certain conditions of differentiability on F and regularity conditions on the "linearized" generalized equation, the mapping G is single-valued, Lipschitz, and "Bouligand" differentiable. Still in the finite-dimensional setting but for general M, King and Rockafellar [2] obtained a kind of generalized derivative for the multifunction G when F is continuous in both variables and strongly Bouligand differentiable in the first, and when a generalized regularity condition is satisfied. Such "differentiability" results are not only very useful for sensitivity analysis, but have application to general approximation schemes as well. However, until now, aside from Robinson's [3] result for a single-valued G, nothing has been said about the "differentiability" of the multifunction G in spaces more general than finite-dimensions, or with the mapping M depending on the parameter p, where this parameter is possibly restricted to a *subset* of a Banach space.

It is clear from equation (1.1.1) that, in general, G will be a *multifunction*, and thus any broad sensitivity theory based on 'differentiation" of such mappings must embrace some notion of generalized derivatives of multifunctions. Rockafellar [5] introduced a generalized derivative, called a "proto-derivative", that applied to multifunctions on Banach spaces. Proto-derivatives are obtained from Painlevé-Kuratowski set limits of the graphs of certain different quotient multifunctions (the graph of a multifunction G is the set of all pairs (x, y) with $y \in G(x)$, and is denoted by $\mathrm{gph}\, G$). In Rockafellar [5, Theorem 5.6], conditions were established that ensured the proto-differentiability of the multifunction $G\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^n$ given by

$$(1.1.2) \qquad G(p, z) = \{x \in C \,|\, z - F(p, x) \in M(x)\}$$

where F is Fréchet differentiable, and M is the normal cone mapping associated with the polyhedral, convex set $C \in \mathbb{R}^n$.

The present paper focuses on a Banach space extension of Rockafellar's [5, Theorem 5.6], where among other things, the parameterization is extended to the mapping M. Specifically, we investigate the "differentiability" of the multifunction G given by

$$(1.1.3) \qquad G(p, z) = \{x \in C \,|\, z - F(p, x) \in M(p, x)\},$$

where p, z, and x are elements of general Banach spaces $\mathscr{P}$, $\mathscr{Z}$, and $\mathscr{X}$. We replace Fréchet differentiability of the mapping F with a form of Hadamard differentiability perhaps only relative to some subset C, called *semi-differentiability relative to C* (when the subset C is actually the whole space, this is simply called *semi-differentiability*). This is a very weak form of differentiability (certainly weaker than Fréchet differentiability), and examples of

semi-differentiable mappings abound (Robinson [3] discussed mappings which have properties stronger than semi-differentiability).

In Theorem 4.1, assuming merely that the single-valued mapping F is semi-differentiable relative to $C_{\mathscr{P}} \times C$ for some set $C_{\mathscr{P}} \subseteq \mathscr{P}$, and the restriction of M to $C_{\mathscr{P}} \times C$, $M_{C_{\mathscr{P}} \times C}$, is proto-differentiable, we show that the (restricted) quasi-solution mapping $G_{C_{\mathscr{P}} \times \mathscr{X}}$ is proto-differentiable, and obtain a formula for its proto-derivative. Furthermore, we do not need the full generality of infinite-dimensional Banach spaces and restricted parameter sets in order to obtain important results from Theorem 4.1. Not only does our Theorem 4.1 generalize Rockafellar's [5, Theorem 5.6] (since the normal cone mapping associated with a polyhedral convex set in $\mathbb{R}^n$ is proto-differentiable), but it provides new sensitivity results for families of *finite-dimensional* nonlinear programming problems with *un*restricted parameterizations (see Example 4.2 and the remark following it).

Besides these finite-dimensional applications, our Theorem 4.1 greatly increases the range of *infinite-dimensional* problems for which the stability of quasi-solution sets can be analyzed. A large and useful class of proto-differentiable mappings in infinite-dimensions has recently been identified. In [6] and [7], the authors showed that for certain composite functions on Hilbert spaces, called "fully amenable" functions (compositions of convex functions with $\mathscr{C}^2$ mappings, and satisfying a constraint qualification), the (Clarke) subgradient mappings of the functions are proto-differentiable as long as the functions themselves are "twice epi-differentiable". Fully amenable functions on finite-dimensional vector spaces have been extensively studied by Poliquin and Rockafellar in [8] and [9]. In the more general Hilbert space setting, Levy [10] showed that many such composite functions are indeed twice epi-differentiable if they happen to be "integral functionals". Such fully amenable integral functionals are common in problems of optimal control, stochastic programming, and the calculus of variations, and thus the subgradient mappings associated with these functionals appear often in first-order conditions for optimality in such problems.

Throughout this paper, $\mathscr{X}$ and $\mathscr{Y}$ will denote general Banach spaces, and we will use strong convergence unless weak convergence is indicated with a " w ". In §2, we investigate several related notions of generalized derivatives of multifunctions, including Rockafellar's "proto-derivative" [5], and in §3 we define semi-differentiability of a single-valued mapping. For locally Lipschitz mappings with domain and range in some finite-dimensional vector spaces, Shapiro [11] showed that semi-differentiability is equivalent to Robinson's "Bouligand" differentiability [12], defined in terms of the "contingent cone". It turns out that proto-differentiation has an important connection with this contingent cone, a fact that suggests there is a relationship between proto-differentiability and semi-differentiability. We establish such a relationship in Proposition 3.4, obtaining the proto-differentiability of the multifunction that is the sum of a proto-differentiable multifunction and a semi-differentiable mapping. Although this result is interesting in its own right, we are currently concerned primarily with its role in the proof of our Theorem 4.1. Section 4 is devoted to the proof of Theorem 4.1, our result giving the proto-differentiability of the parameterized quasi-solution sets (1.1.3). At the end of §4, there are two examples that illustrate some of the uses of Theorem 4.1.

2. Generalized derivatives of multifunctions

An important ingredient in the definition of the proto-derivative of a multifunction is the notion of "graph convergence" which, in turn, depends on a concept of set convergence due to Painlevé and Kuratowski. The *inner set limit as* $t \downarrow 0$, denoted by $\liminf$, of a family of sets $\{G_t\}$ parameterized by $t > 0$, is defined as the set of points x such that for *every* sequence $t_n \downarrow 0$, x is a limit point of some sequence of points taken from the sets G_{t_n}. Likewise, the *outer set limit as* $t \downarrow 0$, denoted by $\limsup$, is defined as the set of points x such that for *some* sequence $t_n \downarrow 0$, x is a limit point of some sequence of points taken from the sets G_{t_n}. In general, the inclusion

$$\liminf_{t \downarrow 0} G_t \subseteq \limsup_{t \downarrow 0} G_t$$

is valid, and "graph convergence" is defined by this inclusion being an equality when the G_t's are the graphs of a family of multifunctions.

Definition 2.1. A family of multifunctions $\{S_t\}$ parameterized by $t > 0$ and mapping $\mathscr{X}$ into $\mathscr{Y}$, graph converges to the multifunction $S: \mathscr{X} \to \mathscr{Y}$ as $t \downarrow 0$, if

$$\limsup_{t \downarrow 0} \operatorname{gph} S_t = \liminf_{t \downarrow 0} \operatorname{gph} S_t = \operatorname{gph} S .$$

The notions of Painlevé-Kuratowski set limits not only help define graph convergence, but have been used to define certain "tangent cones" to sets in Banach spaces. In particular, the *contingent cone to C at a point* $\overline{x} \in C$, is given by

$$T_C(\overline{x}) := \limsup_{t \downarrow 0} \frac{C - x}{t} .$$

For a multifunction Γ from a Banach space $\mathscr{X}$ to another Banach space, a point $x \in \mathscr{X}$, and a point $y \in \Gamma(x)$, Rockafellar [13] defined the *proto-derivative* of Γ at x relative to y as the multifunction whose graph is the limit set of the graphs of the difference quotient multifunctions $(\Delta_t \Gamma)_{x,y}: \mathscr{X} \to Y$, defined by

$$(\Delta_t \Gamma)_{x,y}(\xi) := (1/t)[\Gamma(x + t\xi) - y] \quad \text{for } \xi \in \mathscr{X} \text{ and } t > 0,$$

when these multifunctions happen to graph converge as $t \downarrow 0$.

Definition 2.2. For a multifunction $\Gamma: X \to Y$, a point $x \in X$, and a point $y \in \Gamma(x)$, if the difference quotient multifunctions $(\Delta_t \Gamma)_{x,y}$ graph converge as $t \downarrow 0$, then Γ is *proto-differentiable at x relative to y*. The *proto-derivative* is defined as the multifunction whose graph is the limit set, and is denoted by $\Gamma'_{x,y}$. For any set $C \subseteq \mathscr{X}$, if the restricted multifunction Γ_C, defined by

$$\Gamma_C := \begin{cases} \Gamma(\xi), & \text{if } \xi \in C \\ \varnothing, & \text{otherwise} \end{cases} \quad \text{for } \xi \in \mathscr{X}$$

is proto-differentiable at $x \in C$ relative to $y \in G(x)$, then G is said to be proto-differentiable on C at x relative to y, and the *proto-derivative* is defined as the multifunction whose graph is the limit set, and is denoted by $(\Gamma_C)'_{x,y}$.

Since the notion of proto-differentiation is defined in terms of *graphs* of multifunctions, it is not surprising that the proto-differentiability of a multifunction is equivalent to the proto-differentiability of its inverse multifunction.

Proposition 2.3. *A multifunction* $G: \mathcal{X} \to \mathcal{Y}$ *is proto-differentiable at* x *relative to* $y \in G(x)$ *if and only if its inverse* G^{-1} *is proto-differentiable at* y *relative to* $x \in G^{-1}(y)$, *and the formula* $(G'_{x,y})^{-1} = (G^{-1})'_{y,x}$ *holds.*

Proof. As in [5, Proposition 5.1], this is clear since proto-differentiability is defined in terms of graphs of difference quotient multifunctions. □

In order to state the next lemma, which provides an analytically useful version of the definition of proto-differentiability, we must first introduce a notion of one-sided differentiability of an arc $v: [0, \tau) \to \mathcal{X}$. For such an arc, we define the *one-sided derivative of* v *at* 0 by

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

when this limit exists.

Lemma 2.4. *A multifunction* $G: \mathcal{X} \to \mathcal{Y}$ *is proto-differentiable at* x *relative to* $y \in G(x)$ *with proto-derivative* $G'_{x,y}$ *if and only if for every* $\xi \in \mathcal{X}$,

$$G'_{x,y}(\xi) = \left\{ \eta \, \middle| \, \begin{array}{l} \textit{there exist sequences } \xi_n \to \xi, \, t_n \downarrow 0, \textit{ and} \\ \eta_n \to \eta \textit{ such that } y + t_n \eta_n \in G(x + t_n \xi_n) \end{array} \right\} =: L_+(\xi),$$

and

$$G'_{x,y}(\xi) = \left\{ \eta \, \middle| \, \begin{array}{l} \textit{there exist arcs } v: [0, \tau) \to \mathcal{X} \textit{ and } u: [0, \tau) \to \mathcal{Y} \\ \textit{with } v(0) = x, \, u(0) = y, \, v'_+(0) = \xi, \textit{ and } u'_+(0) = \eta, \\ \textit{and such that } u(t) \in G(v(t)) \textit{ for all } t \in [0, \tau) \end{array} \right\} =: L_-(\xi).$$

Proof. Rockafellar [5, Proposition 2.3] proved this in the case when $\mathcal{X}$ and $\mathcal{Y}$ happen to be finite-dimensional vector spaces, and the same proof works for $\mathcal{X}$ and $\mathcal{Y}$ general Banach spaces. The graphs of the multifunctions L_- and L_+ are, respectively, the lim inf and lim sup of the graphs of the difference quotients for G at (x, y). □

From the definition of the contingent cone, it is evident that the multifunctions L_+ and L_-, as defined in Lemma 2.4, are empty-valued for all $\xi \notin T_{\mathrm{dom}\,G}(x)$ (where $\mathrm{dom}\,G$ denotes the domain of the multifunction G, which is the set of all points $\xi \in \mathcal{X}$ with $G(\xi)$ nonempty). Thus, if the proto-derivative $G'_{x,y}$ exists, its domain is a subset of the contingent cone to the set $\mathrm{dom}\,G$ at the point x.

3. SEMI-DIFFERENTIABILITY

The following notion of semi-differentiability *relative to* a set C allows one to study the differentiability of single-valued mappings with respect to a restricted set of parameters. Parameterized mappings modeling some phenomena might lose their physical relevance for parameters outside a subset C, or could even be undefined outside C. Examples include mappings parameterized by probability measures or by scalars confined to some interval (see Example 4.3). For such mappings, semi-differentiability on the whole space is often inappropriate or even impossible to obtain without artificially extending the mappings to the whole space.

Definition 3.1. Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces, and let C be a subset of $\mathscr{X}$. A mapping $g\colon C \to \mathscr{Y}$ is *directionally semi-differentiable relative to C* at a point $\overline{x} \in \mathscr{X}$ in the direction x, if the *directional semi-derivative relative to C of g at $\overline{x}$ in the direction x*,

$$D_C g_{\overline{x}}(x) := \lim_{\substack{x' \to x \\ t \downarrow 0 \\ \overline{x} + tx' \in C}} \frac{g(\overline{x} + tx') - g(\overline{x})}{t},$$

exists.

A particularly important set of directions is given by the contingent cone to the set C at the point $\overline{x} \in C$, $T_C(\overline{x})$.

Definition 3.2. Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces and let C be a subset of X. A mapping $g\colon C \to \mathscr{Y}$ is *semi-differentiable relative to C* at a point $\overline{x} \in C$, if g is directionally semi-differentiable relative to C at $\overline{x}$ in the direction x for all directions x in the contingent cone to C at $\overline{x}$.

Obviously, any mapping that is semi-differentiable on the whole space $\mathscr{X}$ is necessarily semi-differentiable relative to any subset C. However, for the kinds of mappings on restricted parameter sets that we discussed at the beginning of this section, even though semi-derivatives *relative to a particular subset* might exist, semi-differentiability on the whole space could be impossible to obtain. When the set C is the whole space $\mathscr{X}$, semi-differentiability is the same as Robinson's B-differentiability [3], and is weaker than Fréchet differentiability. In particular, Haraux [14] shows that the projection onto a "polyhedric" set in a Hilbert space in a B-differentiable mapping which is *not* Fréchet differentiable. A polyhedric set is a generalization to Hilbert spaces of a polyhedral set in finite-dimensions (see Mignot [15] for examples of polyhedric sets). Clearly then, supposing semi-differentiability of the mapping F in (1.1.1) is at least as mild as assuming the corresponding differentiability conditions of Robinson [1] or King and Rockafellar [2].

A clue to the relationship between proto-derivatives and Bouligand derivatives can be found in an equivalent definition of the proto-derivative involving the notion of "derivability" of a set.

Definition 3.3. A set C is a Banach space is *derivable* at a point $x \in C$ if the contingent cone to C at x satisfies the following identity:

$$T_C(x) = \liminf_{t \downarrow 0} \frac{C - x}{t}.$$

It is clear from the definition of the contingent cone that Γ is proto-differentiable at x relative to y if and only if the set $\mathrm{gph}\,\Gamma$ is derivable at (x, y). Furthermore, Robinson's original definition of Bouligand differentiability of a *single-valued* mapping f at a point x [12] implies that the graph of f be derivable at $(x, f(x))$. This connection between Bouligand derivatives and proto-derivatives, together with the fact that semi-differentiability is equivalent to Bouligand differentiability under certain circumstances, indicates that the proto-differentiability of a multifunction and the semi-differentiability of a mapping might be related concepts. In fact, a multifunction that is the sum of a proto-differentiable multifunction and a semi-differentiable mapping is itself proto-differentiable.

Proposition 3.4. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces and let C be a subset of $\mathscr{X}$. If the multifunction $H: \mathscr{X} \to \mathscr{Y}$ is empty-valued outside C and proto-differentiable at x relative to $y_0 \in H(x)$, and the mapping $g: C \to \mathscr{Y}$ is semi-differentiable relative to C at x, then the multifunction $G: \mathscr{X} \to \mathscr{Y}$ defined as the sum of H and g, $G := H + g$, is proto-differentiable at x relative to $y := y_0 + g(x) \in G(x)$. Furthermore, the proto-derivative of G is given by: $G'_{x,y} = H'_{x,y_0} + D_C g_x$.*

Proof. Let L_-^G and L_+^G be as in Lemma 2.4 applied to G, and let L_-^H and L_+^H be as in Lemma 2.4 applied to H. For any $\xi \notin T_C(x)$, the sets $L_+^G(\xi)$ and $L_-^G(\xi)$ are both trivially empty (since $\text{dom}\, G \subseteq \text{dom}\, H \subseteq C$, and thus $T_{\text{dom}\, G}(x) \subseteq T_C(x)$). Furthermore, for any $\xi \in T_C(x)$, L_-^H is defined by

$$L_-^H(\xi) = \left\{ \eta \,\middle|\, \begin{array}{l} \text{there exist arcs } v: [0, \tau) \to \mathscr{X} \text{ and } h: [0, \tau) \to \mathscr{Y} \\ \text{with } v(0) = x,\, h(0) = y_0,\, v'_+(0) = \xi, \text{ and } h'_+(0) = \eta, \\ \text{and such that } h(t) \in H(v(t)) \text{ for all } t \in [0, \tau) \end{array} \right\}.$$

By considering arcs $u: [0, \tau) \to \mathscr{Y}$ defined by $u(t) := h(t) + g(v(t))$, we conclude that the inclusion

$$(3.4.1) \qquad L_-^G(\xi) \supseteq L_-^H(\xi) + D_C g_x(\xi)$$

holds. Since H is proto-differentiable at x relative to y_0, Lemma 2.4 implies that $L_-^H = L_+^H$, and inclusion (3.4.1) becomes

$$(3.4.2) \qquad L_-^G(\xi) \supseteq L_+^H(\xi) + D_C g_x(\xi).$$

Clearly, if the inclusion

$$(3.4.3) \qquad L_+^G(\xi) \subseteq L_+^H(\xi) + D_C g_x(\xi)$$

were valid as well, it would follow that $L_+^G(\xi) \subseteq L_-^G(\xi)$. Since the reverse inclusion holds by definition, we could finish the proof by making another appeal to Lemma 2.4. Therefore, we are done if we can show that inclusion (3.4.3) is valid.

Proof of inclusion (3.4.3). By definition, the containment $\eta \in L_+^G(\xi)$ is equivalent to there being sequences ξ_n converging strongly to ξ, η_n converging to η, and $t_n \downarrow 0$ with

$$(3.4.4) \qquad \eta_n \in \frac{H(x + t_n \xi_n) - y_0}{t_n} + \frac{G(x + t_n \xi_n) - g(x)}{t_n}$$

(implicit here is the containment $x + t_n \xi_n \in C$). However, if we define

$$\bar{\eta}_n := \eta_n - \frac{g(x + t_n \xi_n) - g(x)}{t_n},$$

then it follows that $\bar{\eta}_n$ converges to $\eta - D_C g_x(\xi)$, and (3.4.4) implies the containment

$$\bar{\eta}_n \in \frac{H(x + t_n \xi_n) - y_0}{t_n}.$$

This just means that $\eta \in L_+^H(\xi) + D_C g_x(\xi)$, and the claim is proved. $\square$

4. PROTO-DIFFERENTIABILITY OF THE QUASI-SOLUTION SET

We now have all the tools that we need to prove our main theorem, which gives the proto-differentiability of the quasi-solution set.

Theorem 4.1. *Let $\mathcal{P}$, $\mathcal{Z}$ and $\mathcal{X}$ be Banach spaces, and let $C = C_{\mathcal{P}} \times C_{\mathcal{X}}$ with $C_{\mathcal{P}} \subseteq \mathcal{P}$ and $C_{\mathcal{X}} \subseteq \mathcal{X}$. If $G: \mathcal{P} \times \mathcal{Z} \to \mathcal{X}$ is a multifunction of the form*

$$G(p, z) = \{x \in C_{\mathcal{X}} | z - F(p, x) \in M(p, x)\},$$

for a mapping $F: C \to \mathcal{Z}$ which is semi-differentiable relative to C at $(\bar{p}, \bar{x}) \in C$, and a multifunction $M: \mathcal{P} \times \mathcal{X} \to \mathcal{Z}$ which is proto-differentiable on C at $(\bar{p}, \bar{x})$ relative to $\bar{u} := (\bar{z} - F(\bar{p}, \bar{x})) \in M(\bar{p}, \bar{x})$, then G is proto-differentiable on $C_{\mathcal{P}} \times \mathcal{Z}$ at $(\bar{p}, \bar{z})$ relative to $\bar{x}$. Furthermore, the following formula for the proto-derivative of G holds:

$$(G_{C_{\mathcal{P}} \times \mathcal{Z}})'_{(\bar{p}, \bar{z}), \bar{x}}(p, z) = \{x \in \mathcal{X} | z - D_C F_{(\bar{p}, \bar{x})}(p, x) \in (M_C)'_{(\bar{p}, \bar{x}), \bar{u}}(p, x)\}.$$

Proof. In order to prove this theorem, we first define the mappings $\widetilde{M}$ and $\widetilde{F}$ from $\mathcal{P} \times \mathcal{X}$ to $\mathcal{P} \times \mathcal{Z}$, and H, g and A from $\mathcal{P} \times \mathcal{Z}$ to $\mathcal{P} \times \mathcal{X}$ as follows:

$$\widetilde{M}(q, x) := \{0\} \times M(q, x),$$

$$\widetilde{F}(q, x) := (q, F(q, x)),$$

$$H(p, z) := \{(q, x) \in C | (p, z) - \widetilde{F}(q, x) \in \widetilde{M}(q, x)\},$$

$$g(p, z) := (-p, 0),$$

$$A(p, z) := H(p, z) + g(p, z).$$

From the above definitions, it is clear that $A(p, z) = \{(0, x) | x \in G_{C_{\mathcal{P}} \times \mathcal{Z}}(p, z)\}$ and in order to prove our result, we need only show that A is proto-differentiable at $(\bar{p}, \bar{z})$ relative to $(0, \bar{x})$, with the appropriate formula. As a result of Proposition 3.4, this can be accomplished by demonstrating that H is proto-differentiable at $(\bar{p}, \bar{z})$ relative to $(\bar{p}, \bar{x})$ with proto-derivative given by

(4.1.1)
$$\begin{aligned} &H'_{(\bar{p}, \bar{z}), (\bar{p}, \bar{x})}(p, z) \\ &= \{(p, x) \in \mathcal{P} \times \mathcal{X} | z - D_C F_{(\bar{p}, \bar{x})}(p, x) \in (M_C)'_{(\bar{p}, \bar{x}), \bar{u}}(p, x)\}. \end{aligned}$$

Focusing on equation (4.1.1) then as our goal, we notice that the identity

(4.1.2)
$$H^{-1} = \widetilde{F} + \widetilde{M}_C$$

holds. Under our hypotheses, Proposition 3.4 applied to identity (4.1.2) gives the proto-differentiability of H^{-1}, which, by Proposition 2.3, gives the desired proto-differentiability of H. To obtain equation (4.1.1), we first need to calculate the proto-derivative of H^{-1} at $(\bar{p}, \bar{x})$ relative to $(\bar{p}, \bar{z})$. Applying Proposition 3.4 to the identity (4.1.2), we get the following formula for the proto-derivative of H^{-1}:

$$(H^{-1})'_{(\bar{p}, \bar{x}), (\bar{p}, \bar{z})}(p, x) = (p, D_C F_{(\bar{p}, \bar{x})}(p, x) + (M_C)'_{(\bar{p}, \bar{x}), \bar{u}}(p, x)).$$

Therefore, the containment $(q, z) \in (H^{-1})'_{(\bar{p}, \bar{x}), (\bar{p}, \bar{z})}(p, x)$ is equivalent to the conditions $q = p$, and

$$z - D_C F_{(\bar{p}, \bar{x})}(p, x) \in (M_C)'_{(\bar{p}, \bar{x}), \bar{u}}(p, x).$$

Via Proposition 2.3, this is equivalent to equation (4.1.1) and we are done. □

The formula for the proto-derivative of G in Theorem 4.1 can be improved slightly when the sets $C_{\mathscr{P}}$ and $C_{\mathscr{X}}$ happen to be derivable at $\bar{p}$ and $\bar{x}$ respectively. In this case, the identity

$$T_C(\bar{p}, \bar{x}) = T_{C_{\mathscr{P}}}(\bar{p}) \times T_{C_{\mathscr{X}}}(\bar{x})$$

holds, and since the set $(M_C)'_{(\bar{p}, \bar{x}), \bar{u}}(p, x)$ is empty for all $(p, x) \notin T_C(\bar{p}, \bar{x})$, we conclude that the proto-derivative of G actually has the form

$$(G_{C_{\mathscr{P}} \times \mathscr{X}})'_{(\bar{p}, \bar{z}), \bar{x}}(p, z) = \{x \in T_{C_{\mathscr{X}}}(\bar{x}) \mid z - D_C F_{(\bar{p}, \bar{x})}(p, x) \in (M_C)'_{(\bar{p}, \bar{x}), \bar{u}}(p, x)\}.$$

The proof of our Theorem 4.1 follows Rockafellar's proof of [5, Theorem 5.6]. Particular methods that come directly from [5] include our introduction of "extended" mappings at the beginning of the proof and our subsequent reliance on results about the inverse multifunction and the sum of a proto-differentiable multifunction with a differentiable mapping. However, there are some innovations in the proof of our Theorem 4.1, made, for instance, to handle the general multifunction M (recall that in Rockafellar's Theorem 5.6, M was the normal cone mapping associated with a convex, polyhedral set). We call attention to the fact that, just as in Rockafellar [5, Theorem 5.6], the "z"-term must be included in the representation of the mapping G. The presence of this parameter means that the multifunction whose proto-differentiability we are studying, *must at least* depend parametrically on the z and *might* depend, more generally, on other parameters (as represented by the dependence of F and M on p). For the purpose of applying Theorem 4.1 to sensitivity analysis, this "dependence on z" simply means that the parameterized problems whose quasi-solutions are to be studied, must include a linear term as a part of their parametric dependence.

We note finally that Theorem 4.1 holds as well if "strong" is everywhere replaced by "weak". Our choice of convergence is made in order to present both a general statement of the theorem, and one that is applicable to a broad class of problems.

We conclude with two examples that show some of the range of our Theorem 4.1. The first example involves finite-dimensional problems of nonlinear programming.

Example 4.2. Let f_0 be a $\mathscr{C}^2$ function on $\mathbf{R}^d \times \mathbf{R}^n$ and let C be a set defined by finitely many $\mathscr{C}^2$ functions as follows:

$$C = \left\{ x \in \mathbf{R}^n \left| \begin{array}{ll} f_i(x) \leq 0 & \text{for } i = 1, \ldots, s \\ f_i(x) = 0 & \text{for } i = s+1, \ldots, m \end{array} \right. \right\}.$$

For any $p \in \mathbf{R}^d$ and $z \in \mathbf{R}^n$, let $(NLP_{(p,x)})$ denote the problem

$$(NLP_{(p,z)}) \qquad \text{minimize } f_0(p, x) - z \cdot x \quad \text{over } x \in C.$$

A first-order necessary condition for optimality in the problem $(NLP_{(p,z)})$ is given by $-\nabla_x f_0(p, x) + z \in N_C(x)$ (here $N_C(x)$ denotes the (Clarke) normal cone to the set C at the point x), and therefore the solutions to these optimality conditions give a multifunction $G: \mathbf{R}^d \times \mathbf{R}^n \to \mathbf{R}^n$ of the form

$$G(p, z) := \{x \in C \mid z - \nabla_x f_0(p, x) \in N_C(x)\}.$$

If for some point $(\bar{p}, \bar{z}) \in \mathbf{R}^d \times \mathbf{R}^n$, the f_i's defining the set C satisfy the Mangasarian-Fromovitz constraint qualification at a point $\bar{x} \in G(\bar{p}, \bar{z})$, then

the normal cone mapping N_C is proto-differentiable at $\overline{x}$ relative to $\overline{u} := (\overline{z} - \nabla_x f_0(\overline{p}, \overline{x})) \in N_C(\overline{x})$ (see [13]). Clearly, the mapping $\nabla_x f_0 \colon \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ is semi-differentiable at $(\overline{p}, \overline{x})$ relative to the set $\mathbb{R}^d \times C$ (in fact it is $\mathscr{C}^1$ on the whole space), and therefore our Theorem 4.1 gives the proto-differentiability of the quasi-solution mapping G at the point $(\overline{p}, \overline{z})$ relative to $\overline{x}$. The proto-derivative of G has the form

$$G'_{(\overline{p}, \overline{z}), \overline{x}}(p, z) = \{ x \in \mathscr{X} \mid z - \nabla_{x, p} f_0(\overline{p}, \overline{x}) p - \nabla_{x, x} f_0(\overline{p}, \overline{x}) x \in (N_C)'_{\overline{x}, \overline{u}}(x) \}.$$

For this example, the proto-derivative of N_C can be computed using formulas in [13] and [16].

Remark. Many authors have analyzed the sensitivity of solutions to first-order "Lagrangian" necessary conditions for parameterized nonlinear programming problems like $(NLP_{(p, z)})$ (see Kyparisis [17] for a survey). This Lagrangian-based approach produces multipliers that become extra parameters for the problem. Consequently, a direct sensitivity analysis of the solutions to such Lagrangian optimality conditions involves "derivatives" with respect to the multiplier parameters as well as the original parameters. The method in Example 4.2 is based on "primal" necessary conditions, and ultimately this approach also involves multipliers; however they appear only after the proto-differentiation of the quasi-solution mapping, in the formula for the proto-derivative of the normal cone mapping $(N_C)'_{\overline{x}, \overline{u}}$. Thus through our Theorem 4.1, we are able to directly study the sensitivity of the nonlinear programming problem with respect to only the original parameters.

Example 4.3. Let f_0 be a $\mathscr{C}^2$ function on a Hilbert space $\mathscr{X}$ such that for some $\rho > 0$ and $\overline{z} \in \mathscr{X}$, the "$\rho$-coercification" of f_0, $f_0(\cdot) + (\rho/2) \| (\cdot) - \overline{z} \|$, is *convex*, and let $C_{\mathscr{X}}$ be a subset of $\mathscr{X}$. For any real scalar $\lambda \geq \rho$ and any $z \in \mathscr{X}$, let $(P_{(\lambda, z)})$ be the problem of minimizing the λ-coercification of f_0:

$$(P_{(\lambda, z)}) \qquad \text{minimize } f_0(x) + \frac{\lambda}{2} \| x - z \|^2 \quad \text{over } x \in C_{\mathscr{X}}.$$

The quasi-solutions to $(P_{(\lambda, z)})$ are the points $x \in C_{\mathscr{X}}$ that satisfy the optimality condition $\lambda z - \nabla f_0(x) - \lambda x \in N_{C_{\mathscr{X}}}(x)$, and thus the quasi-solution mapping $G \colon \mathbb{R} \times \mathscr{X} \to \mathscr{X}$ has the form

$$G(\lambda, z) := \{ x \in \mathscr{X} \mid z - x - \lambda^{-1} \nabla f_0(x) \in N_{C_{\mathscr{X}}}(x) \}.$$

Using Theorem 4.1, we can study how the solutions to $(P_{(\lambda, z)})$ change as the parameters $\lambda \geq \rho$ and z are perturbed from ρ and $\overline{z}$, by computing the proto-derivative of G *on* $C := [\rho, \infty) \times C_{\mathscr{X}}$ at $(\rho, \overline{z}) \in C$ relative to $\overline{x} \in G(\rho, \overline{z})$. If the normal cone mapping to the set $C_{\mathscr{X}}$ is proto-differentiable at $\overline{x}$ relative to $\overline{u} := (\overline{z} - \overline{x} - \rho^{-1} \nabla f_0(\overline{x}))$, Theorem 4.1 gives the following formula for the proto-derivative of the quasi-solution mapping:

$$(G_C)'_{(\rho, \overline{z}), \overline{x}}(\lambda, z) = \{ x \in \mathscr{X} \mid z + (\lambda/\rho^2) \nabla f_0(\overline{x}) - x - \rho^{-1} \nabla^2 f_0(\overline{x}) x \in (N_{C_{\mathscr{X}}})'_{\overline{x}, \overline{u}} \}.$$

In finite-dimensions, sets like those in Example 4.2 are examples of $C_{\mathscr{X}}$ for which the above analysis is valid (since their associated normal cone mappings are proto-differentiable). In infinite-dimensional Hilbert spaces, sets $C_{\mathscr{X}}$

that can be represented as fully amenable integral functionals also yield proto-differentiable normal cone mappings (see [10] and [6] for details), and therefore families of coercification problems over such sets have proto-differentiable quasi-solution mappings.

REFERENCES

1. S. M. Robinson, *Implicit B-differentiability in generalized equations*, Univ. of Wisconsin-Madison Math. Res. Center, Technical Summary Rep. 2854, 1985.

2. A. J. King and R. T. Rockafellar, *Sensitivity analysis for nonsmooth generalized equations*, preprint, 1991.

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

4. A. L. Dontchev and W. W. Hager, *On Robinson's implicit function theorem*, preprint, 1991.

5. R. T. Rockafellar, *Proto-differentiability of set-valued mappings and its applications in optimization*, Analyse Non Linéaire (H. Attouch, J. P. Aubin, F. H. Clarke, and I. Ekeland, eds.), Gauthier-Villars, Paris, 1989, pp. 449–482.

6. A. B. Levy, R. Poliquin and L. Thibault, *Partial extension of Attouch's theorem with application to proto-derivatives of subgradient mappings*, preprint, 1993.

7. A. B. Levy, *Second-order variational analysis with applications to sensitivity in optimization*, Ph.D. Thesis, University of Washington, 1994.

8. R. A. Poliquin and R. T. Rockafellar, *A calculus of epi-derivatives applicable to optimization*, preprint, 1991.

9. _____, *Amenable functions in optimization*, preprint, 1991.

10. A. B. Levy, *Second-order epi-derivatives of integral functionals*, Set-Valued Analysis **1** (1993).

11. A. Shapiro, *On concepts of directional differentiability*, J. Optim. Theory Appl. **66** (1990), 477–487.

12. S. M. Robinson, *Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity*, Math. Programming Stud. **30** (1987), 45–66.

13. R. T. Rockafellar, *Nonsmooth analysis and parametric optimization*, Methods of Nonconvex Analysis, no. 1446 (A. Cellina, ed.), Springer-Verlag, 1990, pp. 137–151.

14. A. Haraux, *How to differentiate the projection on a convex set in Hilbert space. Some applications to variational analysis*, J. Math. Soc. Japan **29** (1977), 615–631.

15. F. Mignot, *Contrôle dans les inéquations variationelles elliptiques*, J. Funct. Anal. **22** (1976), 130–185.

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

17. J. Kyparisis, *Sensitivity analysis in nonlinear programming under second order assumptions*, Lecture Notes in Control and Information Science, no. 66 (A. V. Balakrishnan and M. Thoma, eds.), Springer-Verlag, New York, 1985, pp. 74–97.

DEPARTMENT OF MATHEMATICS GN-50, UNIVERSITY OF WASHINGTON, SEATTLE, WASHINGTON 98195

E-mail address: levy@math.washington.edu

E-mail address: rtr@math.washington.edu