SENSITIVITY ANALYSIS FOR NONSMOOTH GENERALIZED EQUATIONS

Alan J. King[†] *and R. Tyrrell Rockafellar*[‡]

May 26, 1989; revised March 12, 1991

Abstract. Results pertaining to Lipschitzian and directional differentiability properties for solutions to generalized equations under very general perturbations are obtained with the aid of new differentiation concepts for multivalued maps.

Keywords: sensitivity analysis, generalized equations, contingent derivatives, B-derivatives, implicit function theorems.

[†] IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598, U.S.A.

[‡] Department of Mathematics, University of Washington, Seattle, Washington 98195, U.S.A. Research supported in part by the Air Force Office of Scientific Research under grant AFOSR–89–0081 and the National Science Foundation under grant DMS–8819586.

1. Introduction

Generalized equations are a convenient model for sensitivity analysis in many areas of mathematical programming. There is a considerable body of work concentrating on problems of the form: choose $x \in \mathbb{R}^n$ to satisfy

$$(1.1) \qquad 0 \in f(p, x) + N(x),$$

where f is a given function from $\Omega \times \mathbb{R}^n$ to $\mathbb{R}^m$, N a multifunction from $\mathbb{R}^n$ to $\mathbb{R}^m$, and p an element of an open subset Ω of a normed linear space P. By constructing different multifunctions, one can make this generalized equation cover a wide range of applications from variational inequalities and complementarity problems to first-order necessary conditions in optimization. Our objective in this paper is to develop natural conditions that when holding at a point $\bar{p}$ ensure that the solutions are well-behaved for all p near $\bar{p}$ and are in some sense differentiable at $\bar{p}$, but for a wider class of perturbations than has heretofore been treated in the literature.

An important example of the sort of perturbations we would like to treat in this study, and one that will be covered in a separately derived theorem as a prototype of the sort of analysis we will use on (1.1), is that of a simple generalized equation $0 \in F(x)$ perturbed by the subtraction of a continuous function from the right-hand side, i.e.

$$(1.2) \qquad 0 \in F(x) - p(x).$$

This equation could be molded into the form of (1.1) by defining f to be the evaluation functional, $f(p, x) = p(x)$, in which case the problem should perhaps be studied using the theory developed by Robinson in the papers [**16**] [**17**] [**19**]. However, it seems to fall outside the scope of any of these papers. In [**17**] and [**19**], the theorems guarantee a *unique* solution for all p in a neighborhood of $\bar{p} = 0$. We know this cannot be true of (1.2), since the perturbations, even though continuous, can be wild enough to induce multi-valued solutions no matter how close the perturbations are in sup-norm to zero. (For this reason the results of Dafermos [**2**], Fiacco [**4**], Kyparisis [**11**], or Shapiro [**29**] cannot apply here either.) The result of [**16**] does not impose single-valued solutions, but it does require $f(p, \cdot) = p(\cdot)$ to be Frechet differentiable in x and $\nabla_x f(\cdot, \cdot)$ continuous on a neighborhood of $(\bar{p}, \bar{x})$. Such assumptions are too strong for the study of (1.2).

The perspective developed in this paper permits a study of the perturbed generalized equation (1.1) that is general enough to encompass (1.2) despite its apparent unruly behavior. We first employ a classical fixed point theorem of von Neumann and Kakutani to ensure the existence of solutions, provided that the multifunction $F(\cdot) = f(\bar{p}, \cdot) + N(\cdot)$ satisfies a condition we term *subinvertibility*. We then derive surprisingly strong conclusions concerning derivatives of the solution mapping under only the simple additional assumption that the *contingent derivative* of F^{-1} be everywhere at most a singleton (this allows for the prospect that it may be empty-valued somewhere). No assumptions concerning differentiability in x of the functions $f(p, \cdot)$ are required for any $p \in P$ other than the single point $\bar{p}$.

What we obtain for the sensitivity analysis of (1.1) is not a perfect generalization of the implicit function theorem, because the condition we place on the contingent derivative of F^{-1} is not strong enough to imply subinvertibility of F, unless we impose other conditions as well. For example, when F is *maximal monotone* then single-valuedness of the contingent derivative of F^{-1} at 0 is enough to imply subinvertibility. Taking a different tack, we note that when the multifunction N is *polyhedral* then the contingent derivative of F coincides with the linearization employed by Robinson in his definition of *strong regularity*, and thus under the assumptions of polyhedrality and strong regularity, we show that F is subinvertible and that F^{-1} has a single-valued contingent derivative. This allows us to weaken the differentiability requirements of Robinson's implicit function theorem [**17**], while retaining Lipschitz continuity and B-differentiability of solutions at $\bar{p}$.

A full appreciation of the versatility of the simple formulas we obtain for the derivative of the solution mapping will be attained, we fear, only by the rare reader who has absorbed the implications of the new calculus of set-valued mappings as presented in the papers of Aubin [**1**], Rockafellar [**25**] [**26**] [**27**], and Poliquin [**15**]. We try to convey something of the flavor of this perspective in a brief discussion of second-order sensitivity analysis for mathematical programs.

An important application of this theory arises in the field of stochastic programming and statistical estimation in the process of determining the central limit properties of the sequence of solutions to

$$0 \in \frac{1}{\nu} \sum f(\xi_\nu, x) + N(x), \tag{1.4}$$

where $\{\xi_\nu\}$ is an independent sequence of random observations. For recent treatments of this problem in the stochastic programming literature, see Dupačová and Wets [**3**], King [**6**] [**7**], and Shapiro [**28**]. Under easily satisfied assumptions, it can be shown that the sequence $p^\nu(\cdot) := Ef(\xi_1, \cdot) - \frac{1}{\nu} \sum f(\xi_\nu, \cdot)$ is an asymptotically normal sequence of continuous functions. Our analysis of (1.2) together with the *generalized delta theorem* of [**7**] then gives conditions under which the sequence $\{x^\nu\}$ satisfies a central limit property determined by the asymptotics of the p^ν and the derivative of F. To do justice to this example would require too great an excursion into technical details far from the main theme of the present paper. This will appear in a separate work [**8**].

Earlier versions of the results in this paper are contained in King's dissertation [**6**], where strong monotonicity and Minty's theorem were relied on to provide the required subinvertibility. We are indebted to the associate editor and the referees for a careful and thorough reading—in particular, to one referee for a useful refinement to our definition of subinvertibility—and to Maijian Qian for a valuable discussion of our treatment of second-order sensitivity analysis.

2. Contingent Derivatives and Semidifferentiability.

The notion of derivative of multifunction that we will use is based on the Painlevé–Kuratowski convergence of sets. Letting $\{A_t : t \in T\}$ be a net of sets in a topological space, the limit superior and limit inferior are the sets:

$$\limsup A_t := \{a = \lim_n a_{t_n} \,\big|\, a_{t_n} \in A_{t_n}, n = 1, 2, \ldots\}$$

$$\liminf A_t := \{a = \lim_t a_t \,\big|\, a_t \in A_t, t \in T\}.$$

These set-limits are possibly empty. If they are equal, then the net $\{A_t\}$ has a limit equal to their common value, denoted $\lim_t A_t$.

Let P and X be normed linear spaces, and let H be a multifunction from P into X, that is, $H(p)$ is a subset, possibly empty, of X. Two important sets associated with a multifunction are the *domain*,

$$\operatorname{dom} H := \{p \,\big|\, H(p) \neq \emptyset\},$$

and the *graph*,

$$\operatorname{gph} H := \{(p, x) \,\big|\, x \in H(p)\}.$$

A graphical derivative of H, modelled after the original tangency constructions of Fermat, was recently introduced by Aubin [**1**]: the *contingent derivative* of H at a pair $(\bar{p}, \bar{x}) \in \operatorname{gph} H$ is the mapping, which we shall denote $DH(\bar{p}|\bar{x})$, whose graph is the *contingent cone* to the graph of H at $(\bar{p}, \bar{x})$. This is summarized in the formula

$$\operatorname{gph} DH(\bar{p}|\bar{x}) = \limsup_{t \downarrow 0} t^{-1}[\operatorname{gph} H - (\bar{p}, \bar{x})]. \tag{2.1}$$

The contingent derivative always exists, and its graph is a cone that includes the origin in $P \times X$. The contingent derivative of the inverse of H is just the inverse of the contingent derivative, and we will denote it $DH^{-1}(\bar{x}|\bar{p})$.

This perspective was extended by Rockafellar [**26**] in two directions, each a tightening of the limiting behavior in (2.1). If the limit exists in (2.1), then we say that H is *proto-differentiable* at $(\bar{p}, \bar{x})$. Proto-differentiation of subgradient mappings is related to second-order epi-differentiability of objective functions, on which Rockafellar [**25**] [**27**] bases a comprehensive study of necessity and sufficiency of second-order optimality conditions. A stronger property is *semi-differentiability*, which requires that the limit

$$\lim_{\substack{t \downarrow 0 \\ w' \to w}} (H(\bar{p} + tw') - \bar{x})/t$$

exist for all w. When it does, it equals the contingent derivative $DH(\bar{p}|\bar{x})(w)$. Semi-differentiability is closer to "Hadamard" directional differentiability, in the sense that the derivative should not depend on how w' approaches w. Because of this uniformity, semi-differentiability is an extremely useful property in many applications (as discussed in the introduction, for example).

King [**6**] showed that the single-valuedness of the contingent derivative acts like a regularity condition that can be used to prove the semi-differentiability of solution mappings to generalized equations. The next two propositions elaborate this idea, and provide the platform on which the sensitivity analysis of the generalized equations (1.1) and (1.2) will be erected.

We recall the following definition from [**16**]. A multifunction H from a normed linear space P into a normed linear space X is *upper Lipschitzian* at $\bar{p}$ if there are a number $\lambda \geq 0$ and a neighborhood Ω of $\bar{p}$ with

$$H(p) \subset H(\bar{p}) + \lambda \|p - \bar{p}\| B$$

for all $p \in \Omega$, where here, as elsewhere in this paper, the set B is the unit ball of the space in question. It should be noted that single-valued mappings that are upper Lipschitzian are not necessarily locally Lipschitz continuous, but the converse always holds. In these next propositions, we suppose that X is locally compact, i.e. $X = \mathbb{R}^n$.

Proposition 2.1. *Let $H : P \rightrightarrows \mathbb{R}^n$ be such that at a point $\bar{p}$ and a point $\bar{x} \in H(\bar{p})$ the set $DH(\bar{p}|\bar{x})(0)$ contains at most the single element 0. Then there is a neighborhood U of $\bar{x}$ such that $U \cap H(\bar{p}) = \{\bar{x}\}$, and $U \cap H$ is upper Lipschitzian at $\bar{p}$.*

Proof. Apply arguments similar to [**26, Theorem 4.1**]. The idea is that if either of the conclusions fails to hold then one can easily construct a nonzero element of $DH(\bar{p}|\bar{x})$. ☐

Proposition 2.2. *Let $H : P \rightrightarrows \mathbb{R}^n$ be such that at a point $\bar{p}$ and a point $\bar{x} \in H(\bar{p})$ one has for every $w \in P$ the set $DH(\bar{p}|\bar{x})(w)$ contains at most a single element, and for* *every neighborhood U' of $\bar{x}$ the set $H^{-1}(U')$ is a neighborhood of $\bar{p}$. Then, in addition to satisfying the conclusions of Proposition 2.1, H is semi-differentiable at $\bar{p}$ relative to $\bar{x}$, and for every $w \in P$ one has*

$$DH(\bar{p}|\bar{x})(w) = \limsup_{t \downarrow 0}[H(\bar{p} + tw) - \bar{x}]/t.$$

In particular, $DH(\bar{p}|\bar{x}) : P \to \mathbb{R}^n$ is a continuous function that is Lipschitz at 0. Furthermore, every selection $x(p) \in H(p) \cap U$ is upper Lipschitzian and semi-differentiable at $\bar{p}$ with $Dx(\bar{p}|\bar{x}) = DH(\bar{p}|\bar{x})$.

Proof. Our single-valuedness assumption implies in particular that the only element of $DH(\bar{p}|\bar{x})(0)$ is the single element 0, so Proposition 2.1 applies. Now let D_t be the difference quotient multifunction

$$D_t(w) = t^{-1}[U \cap H(\bar{p} + tw) - \bar{x}], \qquad t > 0.$$

We want to show that

$$\lim_{\substack{t \downarrow 0 \\ w' \to w}} D_t(w') = DH(\bar{p}|\bar{x})(w).$$

Since $\bar{p}$ is in the interior of $H^{-1}(U)$ by assumption, it follows that $D_{t^\nu}(w^\nu)$ is eventually nonempty for any sequence $t^\nu \downarrow 0$ and $w^\nu \to w$. The upper Lipschitzian property of $U \cap H$ implies that eventually

$$D_{t^\nu}(w^\nu) \subset \lambda \|w\| B.$$

Thus any sequence $u^\nu \in D_{t^\nu}(w^\nu)$ is eventually bounded and has cluster points. These cluster points must be in $DH(\bar{p}|\bar{x})(w)$, by definition of the contingent derivative, thus $DH(\bar{p}|\bar{x})(w)$ is nonempty. Since, by our assumption, $DH(\bar{p}|\bar{x})(w)$ is at most a singleton, say $\{u\}$, we have in fact shown that $\lim u^\nu = u$ for all sequences $u^\nu \in D_{t^\nu}(w^\nu)$, all $t^\nu \downarrow 0$, and all $w^\nu \to w$; so H is semi-differentiable at $\bar{p}$ relative to $\bar{x}$. Since $DH(\bar{p}|\bar{x})$ is everywhere nonempty and single-valued, and has closed graph, it is therefore a continuous function. $DH(\bar{p}|\bar{x})$ is Lipschitz at 0 because H is upper Lipschitzian at $\bar{x}$. The final statement concerning selections is trivial. □

The reader may wonder whether the assumptions of Proposition 2.2 are sufficient to imply that H itself is single-valued near $\bar{p}$. This is false, even for maximal monotone multifunctions, as the following example shows. For p greater than $\pi/4$ let $H(p) = 1$. Set $p_1 = \pi/4$ and let $H(p_1)$ equal the closed interval $[\sin p_1, \tan p_1]$. Extend the graph of H to the left of p_1, with value identically equal to $\sin p_1$, until it touches the graph of the tangent curve, at the point $p_2 = \tan^{-1}(\sin p_1)$, and set $H(p_2) = [\sin p_2, \tan p_2]$. Continue in this way; the sequence $\{p_\nu\}$ converges to 0. Repeat the process, but with directions reversed, beginning from $-\pi/4$. The multifunction H so constructed is a maximal monotone multifunction satisfying the conditions of Proposition 2.2 (it is semi-differentiable at 0 with $DH(0|0)(w) = w$), but H is multiple-valued in any neighborhood of 0.

In preparation for the results of the next few sections, let us briefly consider the connection between contingent derivatives and directional derivatives of functions. If a continuous function $h : P \to \mathbb{R}^n$ has a contingent derivative at $\bar{p}$ that is single-valued everywhere, then Proposition 2.2 states that the limit

$$\lim_{\substack{t\downarrow 0\\ w'\to w}} \frac{h(\bar{p}+tw') - h(\bar{p})}{t}$$

exists for all w. This property has received a lot of attention in the literature, but is known under a variety of names. Rockafellar [**24**] says that h is directionally differentiable *in the Hadamard sense* at $\bar{p}$, while Robinson [**20**] calls it *Bouligand differentiability* (B-differentiability, for short) in honor of the man who introduced the contingent cone. Since the latter definition is closer in spirit to our perspective here, we adopt Robinson's terminology. When P is finite-dimensional, it is well known that if h is directionally differentiable in the *ordinary sense*, i.e.

$$\lim_{t\downarrow 0} \frac{h(\bar{p}+tw) - h(\bar{p})}{t} = Dh(\bar{p})(u),$$

and $Dh(\bar{p})(\cdot)$ is continuous, then h is also B-differentiable. See Shapiro [**30**] for a review of the equivalences between various definitions of directional differentiability.

3. The Basic Case.

To ensure that solutions exist for the perturbed generalized equation (1.2), we formulate a condition that requires that the graph of F contains a subset that can be regarded as the graph of a multifunction to which the von Neumann–Kakutani fixed point theorem can be applied. (A similar condition appears in Kummer's papers [**9**] and [**10**].) We say that a multifunction $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *subinvertible* at $(\bar{x}, 0)$, if one has $0 \in F(\bar{x})$ and there exist a compact convex neighborhood U of $\bar{x}$ in $\mathbb{R}^n$, a positive constant $\varepsilon > 0$, and a nonempty convex-valued mapping $G : \varepsilon B \rightrightarrows U$ such that: $\operatorname{gph} G$ is closed, the point $\bar{x}$ belongs to $G(0)$, and $G(y)$ is contained in $F^{-1}(y)$ for all $y \in \varepsilon B$. For instance, F is subinvertible at $(\bar{x}, 0)$ if there exists a selection $x(y)$ of $F^{-1}(y)$ that is continuous on a compact neighborhood of 0, with $x(0) = \bar{x}$.

For convenience of presentation, we call a multifunction $Q : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ a *perturbation multifunction* on a set U in $\mathbb{R}^n$ if $\operatorname{gph} Q$ is a closed subset of $U \times \mathbb{R}^n$ and for every $x \in U$ the set $Q(x)$ is nonempty and convex. The *width* of Q relative to a bounded set $U \subset \mathbb{R}^n$ is defined to be the quantity

$$\|Q\|_U = \sup_{x \in U} \sup_{y \in Q(x)} |y|.$$

The definitions lead to the following proposition. In our applications, the perturbation multifunctions will be continuous functions.

Proposition 3.1. *Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a given multifunction that is subinvertible at $(\bar{x}, 0)$ in $\mathbb{R}^n$. Then there are a compact convex subset $U \subset \mathbb{R}^n$ and a real number $\varepsilon > 0$ such that for the solution mapping*

$$J(Q) = \{x \in \mathbb{R}^n \mid 0 \in F(x) - Q(x)\}$$

one has that $U \cap J(Q)$ is nonempty for every perturbation multifunction $Q : U \rightrightarrows \mathbb{R}^m$ satisfying $\|Q\|_U \leq \varepsilon$. If, moreover, the mapping $U \cap F^{-1}$ is U.L.(λ) at 0, then

$$U \cap J(Q) \subset J(0) + \lambda \|Q\|_U B$$

for all perturbation multifunctions Q with width $\|Q\|_U$ sufficiently close to 0.

Proof. Let a compact convex neighborhood U of $\bar{x}$, constant $\varepsilon > 0$, and multifunction $G : \varepsilon B \rightrightarrows U$ be such as are guaranteed by the subinvertibility assumption on F, and let $Q : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be any perturbation multifunction with $\|Q\|_U$ less than ε. Apply the von Neumann–Kakutani fixed point theorem [**5, Theorem 2**] to the sets $\operatorname{gph} G$ and $\operatorname{gph} Q$ and conclude that there is at least one fixed point $x_Q \in G(Q(x_Q))$. Since $G(y) \subset U \cap F^{-1}(y)$ for all $y \in \varepsilon B$, it follows that $x_Q \in U \cap J(Q)$. This proves existence. The Lipschitz condition for $U \cap J$ can easily be derived from the observation that

$$U \cap J(Q) \subset U \cap F^{-1}(y), \quad \forall y \in \varepsilon B$$

and the proof is complete. □

The two conditions, single-valuedness of the contingent derivative of F^{-1} and subinvertibility of F, combine to produce the main result of this section. In this theorem, and elsewhere in the paper, we denote the space of continuous functions from a set U to the space $\mathbb{R}^m$ as $\mathcal{C}_m(U)$.

Theorem 3.2. *In the generalized equation (1.2), let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a multifunction that is subinvertible at $(\bar{x}, 0)$. Then the following statements apply to this generalized equation:*
(a) If $DF^{-1}(0|\bar{x})(0)$ contains the single element 0, then there exists a neighborhood U of $\bar{x}$ in $\mathbb{R}^n$ such that for the solution mapping $J : \mathcal{C}_m(U) \rightrightarrows \mathbb{R}^n$ defined by

$$J(p) = \{x \in \mathbb{R}^n \,\big|\, 0 \in F(x) - p(x)\}$$

one has: $U \cap J(0) = \{\bar{x}\}$, $U \cap J$ is upper Lipschitzian at 0, and $U \cap J(p)$ is nonempty for every p in a neighborhood of 0 in $\mathcal{C}_m(U)$.
(b) If for every $y \in \mathbb{R}^m$ the set $DF^{-1}(0|\bar{x})(y)$ is at most a singleton, then J is semi-differentiable at $(0, \bar{x})$ with derivative given by

$$DJ(0|\bar{x})(w) = \{u \in \mathbb{R}^n \,\big|\, 0 \in DF(\bar{x}|0)(u) - w(\bar{x})\},$$

i.e. $DJ(0|\bar{x})(w) = DF^{-1}(0|\bar{x})(w(\bar{x}))$. Furthermore, all selections $x(p) \in J(p) \cap U$ are upper Lipschitzian and B-differentiable at 0, with $Dx(0) = DJ(0|\bar{x})$.

Proof. By Proposition 2.1 applied to F^{-1}, there is a neighborhood U_1 of $\bar{x}$ such that $U_1 \cap F^{-1}(0) = \{\bar{x}\}$ and $U_1 \cap F^{-1}$ is upper Lipschitzian at 0 in $\mathbb{R}^m$. The subinvertibility of F at $(\bar{x}, 0)$ and the upper Lipschitzian property of $U_1 \cap F^{-1}$ at 0 together imply that the restriction of F to U_1 is subinvertible at $(\bar{x}, 0)$, with the mapping $U_1 \cap G$ serving as that required by the definition of subinvertibility, where G is the mapping guaranteed by the subinvertibility of F. To see this, note that the only thing that could go wrong is that there could be a sequence $\{y^\nu\}$ converging to 0 in $\mathbb{R}^m$ such that $G(y^\nu) \cap U_1 = \emptyset$ for all ν. Since G is nonempty-valued and has closed graph, there must exist points of $G(0)$ outside of U_1. But $G(0)$ is a convex set containing $\bar{x}$, and furthermore, $G(0)$ is contained in $F^{-1}(0)$. This contradicts the single-valuedness of $U_1 \cap F^{-1}(0)$; hence, F restricted to U_1 must be subinvertible at $(\bar{x}, 0)$. Now apply Proposition 3.1 and conclude that there is a neighborhood $U \subset U_1$ of $\bar{x}$ such that $U \cap J(p)$ is nonempty for every p sufficiently near 0 in $\mathcal{C}_m(U)$ and $U \cap J$ is upper Lipschitzian at 0 in $\mathcal{C}_m(U)$. This argument proves statement (a). It also proves that for every neighborhood $U' \subset U$ of $\bar{x}$ in $\mathbb{R}^n$ the set $(J)^{-1}(U')$ is a neighborhood of 0 in $\mathcal{C}_m(U)$. By Proposition 2.2 applied to J and by the single-valuedness of $DF^{-1}(0|\bar{x})$, to prove statement (b) it remains only to show that the contingent derivative of J satisfies $DJ(0|\bar{x})(w) \subset DF^{-1}(0|\bar{x})(w(\bar{x}))$ for all $w \in \mathcal{C}_m(U)$. So, let $u \in DJ(0|\bar{x})(w)$. Then there must exist positive numbers $t^\nu \to 0$, n-vectors $u^\nu \to u$, and $\mathbb{R}^m$-valued continuous functions $w^\nu \to w$ (uniformly on U), with

$$w^\nu(\bar{x} + t^\nu u^\nu) \in F(\bar{x} + t^\nu u^\nu)/t^\nu. \tag{3.2}$$

Since $w^\nu(\bar{x}+t^\nu u^\nu) \to w(\bar{x})$, it follows from the definition of the contingent derivative that $w(\bar{x}) \in DF(\bar{x}|0)(u)$, which is what we set out to show. □

This is a surprising result. It indicates that, in the situation described in the theorem, the result of differentiating with respect to continuous perturbations is identical to that of differentiating with respect to constant perturbations! As an illustration, consider the trivial generalized equation

$$0 \in Ix,$$

where I is the identity mapping on $\mathbb{R}$. This equation obviously fulfills all of the conditions of the theorem; hence the solution mapping

$$J(p) = \{x \in \mathbb{R} \,|\, 0 \in x - p(x)\}$$

has a single-valued semi-derivative with respect to perturbations in $\mathcal{C}(U)$ for any neighborhood U of 0, and

$$DJ(0|0)(w) = \{u \,|\, 0 \in u - w(0)\} = w(0).$$

Now we know that every neighborhood of the identity mapping in $\mathcal{C}(U)$ contains continuous functions that are as wild as an analyst's nightmare, so $J(p)$ must be multiple-valued at some p in every neighborhood of 0 in $\mathcal{C}(U)$. But, how can a multivalued mapping have a derivative that is everywhere single-valued? Because we are restricted to *continuous* perturbations. In the present example, recall that by (3.2) a point $(w(\cdot), u)$ lies in the graph of $DJ(0|0)$ if and only if there is a net of points $u^t \to u$ such that

$$0 \in 0 + tu^t - tw(0 + tu^t), \quad t > 0.$$

So $u^t = w(tu^t)$ and since w is continuous at 0 there must by definition be only one cluster point as $t \to 0$.

4. Sensitivity Analysis for Nonsmooth Generalized Equations.

With Theorem 3.3 as a prototype, we study in this section the sensitivity of the perturbed generalized equation (1.1). A critical step in Theorem 3.3 is the convergence of the sequence on the left side of the inclusion (3.2). This corresponds in the setting of (1.1) to ensuring the existence of unique limits for sequences of the form

$$[f(\bar{p}+t^\nu w^\nu, x^\nu) - f(\bar{p}, x^\nu)]/t^\nu, \tag{4.1}$$

for any sequence of positive numbers $\{t^\nu\}$ converging to zero, any sequence $\{w^\nu\}$ converging to w in P, and any sequence of vectors $\{x^\nu\}$ converging, no matter how slowly, to $\bar{x}$. Clearly, it is necessary that $f(\cdot,\bar{x})$ be B-differentiable at $\bar{p}$. In this case we say that f has a *partial* B-derivative in p at $(\bar{p},\bar{x})$ and denote this B-derivative by $D_p f(\bar{p},\bar{x})$. But it is also necessary that (4.1) be uniformly close to $D_p f(\bar{p},\bar{x})$. We will see the required uniformity is a consequence of the following definition of Robinson [**21**]: the partial B-derivative $D_p f(\bar{p},\bar{x})$ is *strong* if for each $\varepsilon > 0$ there exist neighborhoods Ω of 0 in P and U of $\bar{x}$ in $\mathbb{R}^n$ such that for every $x \in U$ the function

$$w \mapsto f(\bar{p}+w, x) - f(\bar{p}, x) - D_p f(\bar{p},\bar{x})(w)$$

is continuous with Lipschitz constant ε on Ω.

We are now ready to formulate the main result of this paper. In this theorem we show that the imposition of subinvertibility and single-valuedness of the inverse of the contingent derivative for the single multifunction $F(\cdot) = f(\bar{p},\cdot) + N(\cdot)$ is sufficient to imply semi-differentiability of the solution mapping at $\bar{p}$. The only other assumptions required are joint continuity of f and the existence of a strong partial B-derivative of f in p at $(\bar{p},\bar{x})$.

Theorem 4.1. *In the generalized equation (1.1), let Ω be an open subset of a normed linear space P, let $N : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a multifunction, and let $f : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ be a continuous function that has a strong partial B-derivative in p at the pair $(\bar{p},\bar{x})$, where $\bar{x}$ is a solution to (1.1) at $\bar{p}$. Define the mapping $F(x) = f(\bar{p},x) + N(x)$, and suppose that F is subinvertible at $(\bar{x},0)$. Then the following statements apply to this generalized equation:*
(a) If $DF^{-1}(0|\bar{x})(0)$ contains the single element 0, then there is a neighborhood U of $\bar{x}$ such that for the solution mapping

$$J(p) = \{x \in \mathbb{R}^n \,\big|\, 0 \in f(p,x) + N(x)\}$$

one has: $U \cap J(\bar{p})$ is the singleton $\{\bar{x}\}$, $U \cap J$ is upper Lipschitzian at $\bar{p}$, and for all p in a neighborhood of $\bar{p}$ the set $U \cap J(p)$ is nonempty.

(b) If for every $y \in \mathbb{R}^m$ the set $DF^{-1}(0|\bar{x})(y)$ consists of at most a single element, then J is semi-differentiable at $\bar{p}$ and

$$DJ(\bar{p}|\bar{x})(w) = \{u \in \mathbb{R}^n \mid 0 \in D_p f(\bar{p}, \bar{x})(w) + DF(\bar{x}|0)(u)\}$$

i.e. $DJ(\bar{p}|\bar{x})(w) = DF^{-1}(0|\bar{x})(-D_p f(\bar{p}, \bar{x})(w))$. Furthermore, every selection $x(p) \in U \cap J(p)$ is upper Lipschitzian and B-differentiable at $\bar{p}$, with $Dx(\bar{p}) = DJ(\bar{p}|\bar{x})$.

Proof. For statement (a), apply the argument of the first part of Theorem 3.3 to F with perturbations $f(\bar{p}, \cdot) - f(p, \cdot)$. The continuity of f assures that the supremum of $|f(\bar{p}, x) - f(p, x)|$ can be made small uniformly in x for all p sufficiently near $\bar{p}$. For statement (b), we proceed similarly as in Theorem 3.3: we prove that $DJ(\bar{p}|\bar{x})(w)$ is equal to $DF^{-1}(0|\bar{x})(-D_p f(\bar{p}, \bar{x})(w))$, and the rest follows from Proposition 2.2 and the single-valuedness of $DF^{-1}(0|\bar{x})$. Let $u \in DJ(\bar{p}|\bar{x})(w)$ for some $w \in P$. Then there must exist sequences $\{t^\nu\}$ converging to 0, $\{w^\nu\}$ converging to w in P, and $\{u^\nu\}$ converging to u in $\mathbb{R}^n$ such that

$$0 \in F(\bar{x} + t^\nu u^\nu) - [f(\bar{p}, \bar{x} + t^\nu u^\nu) - f(\bar{p} + t^\nu w^\nu, \bar{x} + t^\nu u^\nu)]$$

for all ν. Let $\varepsilon > 0$. Since $D_p f(\bar{p}, \bar{x})$ is strong, it follows that for all ν sufficiently large one has

$$-D_p f(\bar{p}, \bar{x})(w^\nu) \in F(\bar{x} + t^\nu u^\nu)/t^\nu + \varepsilon \|w^\nu\| B,$$

where B is the unit ball in $\mathbb{R}^m$. Letting $\nu \to \infty$ and observing that ε was arbitrary, finishes the proof. □

Remark 4.2. Theorem 4.1 applies to the situation covered in Theorem 3.3 by taking f equal to the evaluation map $f(p, x) = p(x)$. It is easy to see that f is strongly B-differentiable in the first variable, with $D_p f(p, x)(w) = w(x)$, so Theorem 3.3 follows from 4.1. This example generalizes to a large class of mappings. Let Q be a Banach space and Q^* the dual space of continuous linear functionals on Q. Define two functions $r : P \to Q$ and $s^* : \mathbb{R}^n \to Q^*$, and let

$$f(p, x) = \langle s^*(x), r(p) \rangle.$$

If the function r is B-differentiable at a point $\bar{p}$ in P, then f is partially B-differentiable in p at $(\bar{p}, \bar{x})$ for any $\bar{x}$ in $\mathbb{R}^n$ with

$$D_p f(\bar{p}, \bar{x})(w) = \langle s^*(x), D_p r(\bar{p})(w) \rangle.$$

Furthermore, this derivative is strong if $D_p r(\bar{p})$ is strong and $x \mapsto \|s^*(x) - s^*(\bar{x})\|$ is continuous at $\bar{x}$. In the case of the evaluation functional, r is the identity map on $Q = \mathcal{C}_m(U)$ and $s^*(\bar{x})$ the element of Q^* that evaluates $q \in Q$ at $\bar{x}$.

Remark 4.3. By making further assumptions on f and N we can state a convenient form for the derivative DF. Robinson [**21**] showed that if f is separately B-differentiable in both variables at $(\bar{p}, \bar{x})$ and one of these partial B-derivatives is strong, then f is B-differentiable as a function of both variables and we have the addition formula

$$Df(\bar{p}, \bar{x})(w, u) = D_p f(\bar{p}, \bar{x})(w) + D_x f(\bar{p}, \bar{x})(u).$$

Furthermore, Rockafellar [**26**] showed, in effect, that if f is partially B-differentiable in x at $(\bar{p}, \bar{x})$ and N is proto-differentiable at $(\bar{x}, -f(\bar{p}, \bar{x}))$, then the multifunction F given by $F(x) = f(\bar{p}, x) + N(x)$ is actually proto-differentiable at $(\bar{x}, 0)$ with

$$DF(\bar{x}|0)(u) = D_x f(\bar{p}, \bar{x})(u) + DN(\bar{x}| - f(\bar{p}, \bar{x}))(u).$$

This formula can be substituted for DF in Theorem 4.1 whenever $f(\bar{p}, \cdot)$ is B-differentiable at $\bar{x}$ and N is proto-differentiable at $(\bar{x}, -f(\bar{p}, \bar{x}))$.

Applying the above formula for DF in the case when P is finite-dimensional, f is jointly differentiable at $(\bar{p}, \bar{x})$, and N is identically 0, we see that Theorem 4.1 contains the statement of the classical implicit function theorem:

$$\nabla_p J = -(\nabla_x f)^{-1} \nabla_p f.$$

But the comparison illustrates an important shortcoming of our theorem. In the classical result, the assumption that $(\nabla_x f)^{-1}$ is single-valued at 0 *implies* the single-valuedness of $(\nabla_x f)^{-1}$ and the invertibility of f. In our theorem, the single-valuedness of $DF^{-1}(0|\bar{x})$ at zero, the single-valuedness of $DF^{-1}(0|\bar{x})$, and the subinvertibility of F are all distinct notions. The source of the difficulty is the generality of the contingent derivative: there are many multifunctions that are not subinvertible, but that have deceptively well-behaved contingent derivatives. New research points to an implicit function theorem that would involve conditions only on a certain *adjoint derivative*, replacing the assumption of subinvertibility and putting our theorem on an even footing with the classical statement. The research on this topic is not ripe for publication at this writing; however, when F is maximal monotone or N has the special form of a polyhedral multifunction, then much more can be said.

5. Application to Maximal Monotone Multifunctions.

Maximal monotone multifunctions play an important role in the theory of generalized equations, in the sense that one expects the best possible behavior from this class. And we shall not be disappointed here either, since it turns out that we are able to determine a very precise relationship between subinvertibility and single-valuedness of the inverse of the contingent derivative at 0. For the definition of maximal monotonicity and important motivating examples, we refer the reader to [**23**].

Theorem 5.1. *Let* $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *be a maximal monotone multifunction, and suppose that at a point* $\bar{x}$ *in* $\mathbb{R}^n$ *with* $0 \in F(\bar{x})$ *one has that the only element of* $DF^{-1}(0|\bar{x})(0)$ *is* 0 *itself. Then, in addition to all the conclusions of Proposition 2.1,* F *is subinvertible at* $(\bar{x}, 0)$.

Proof. We will show that the inverse F^{-1} is a closed, locally bounded, nonempty and convex-valued multifunction, and thus fulfills the requirements of the definition of subinvertibility. Since F is maximal monotone, so is F^{-1}; hence F^{-1} is convex-valued and closed. It remains to show that F^{-1} is nonempty and locally bounded. An application of Proposition 2.1 establishes that the set $F^{-1}(0)$ must consist of $\bar{x}$ alone, and this in turn implies that 0 is in the interior of the domain D of F^{-1}, as can be seen by the following argument. The closure of the domain D is a convex set, by Minty [**13**], hence if 0 is not in the interior of D we can add to $F^{-1}(0)$ all the points in the normal cone to $\operatorname{cl} D$ at 0 without destroying the monotonicity of F^{-1}. By maximality, all such points must already belong to $F^{-1}(0)$, contradicting the proven single-valuedness of F^{-1} at 0. It follows that F^{-1} is nonempty-valued on a neighborhood of 0. To finish the proof, we apply a theorem of Rockafellar [**22**] which says that a maximal monotone multifunction is locally bounded on the interior of its domain. □

As a corollary to this theorem, one derives directly from Theorem 4.1 an implicit function theorem for the generalized equation (1.1) when the perturbed multifunction F is maximal monotone.

Corollary 5.2. *For the generalized equation (1.1), let all the conditions of the preamble in Theorem 4.1 hold (with* $m = n$*), except that instead of supposing that* F *is subinvertible at* $(\bar{x}, 0)$*, suppose only that* F *is maximal monotone. Then the statements (a) and (b) of Theorem 4.1 apply to this generalized equation.*

For a maximal monotone multifunction F, is the single-valuedness of $DF^{-1}(0|\bar{x})$ at 0 sufficient to imply the single-valuedness of $DF^{-1}(0|\bar{x})$ at all points of $\mathbb{R}^n$? The following example shows that this is not true. Let F be the subdifferential of the convex function

$$h(x_1, x_2) = \frac{1}{2}(|x_1| + |x_2|)^2.$$

Then

$$F(x_1, x_2) = [|x_1| + |x_2|]\,(g(x_1) \times g(x_2))$$

where g is the subdifferential of the absolute value function on $\mathbb{R}$. Moreover, F equals its contingent derivative at $(0,0)$; that is, $F = DF(0|0)$ since $F(tx_1, tx_2) = tF(x_1, x_2)$ for all $t \geq 0$. Thus, we have indeed $DF^{-1}(0|0)(0) = \{0\}$, but for example, for all $t > 0$ we have

$$DF^{-1}(0|0)(t,t) = \{(x_1, x_2) \,|\, x_1 \geq 0, x_2 \geq 0, \text{ and } x_1 + x_2 = t\},$$

which is multiple-valued.

6. Polyhedral Multifunctions and Strong Regularity.

Many particular properties of mathematical programs and other applications modelled by the generalized equation (1.1) may be attributed to the fact that the multifunction $N : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a *polyhedral multifunction*, that is, one whose graph is the union of finitely many convex polyhedra in $\mathbb{R}^n \times \mathbb{R}^m$. For the uses of such multifunctions and a review of their many pleasant properties, we refer the reader to Robinson [**16**] [**18**] [**19**]. More specifically, the first-order optimality conditions for optimization problems whose objective can be represented by a smooth function composed with a piecewise linear-quadratic convex function can be placed in the form (1.1) with polyhedral N, provided a certain basic constraint qualification is satisfied; cf. Rockafellar [**25**] [**27**] and Poliquin [**15**].

The general strategy reflected in the theorems of this section is to apply readily available implicit function theory to the generalized equation

$$0 \in f(\bar{p}, x) - y + N(x), \tag{6.1}$$

which may be viewed as a generalized equation perturbed by the subtraction of y, in order to verify, especially, the subinvertibility of the multifunction $F(\cdot) = f(\bar{p}, \cdot) + N(\cdot)$. With the subinvertibility then assured, an appeal to Theorem 4.1 gives the desired behavior of the solution mapping under only the additional and easily verified assumption of the single-valuedness of the contingent derivative of the inverse of F. Our demonstration of this technique employs the implicit function theory of Robinson in [**17**] and [**21**].

We give two results: one for general polyhedral multifunctions but differentiable f, and one for maximal monotone polyhedral multifunctions and B-differentiable f. The key property of polyhedral multifunctions that makes both theorems work is the trivial fact (but one whose proof is unavoidably long) that a polyhedral multifunction N is proto-differentiable at every pair $(\bar{x}, \bar{y})$ in its graph, and to every such pair there correspond neighborhoods U and V, respectively, such that

$$\text{gph}\, N \cap (U \times V) = [\text{gph}\, DN(\bar{x}|\bar{y}) + (\bar{x}, \bar{y})] \cap (U \times V). \tag{6.2}$$

This fact allows us to link contingent derivatives with the linearization employed by Robinson in the proof of the implicit function theorem in [**17**]. Assume that f is partially B-differentiable in x at $(\bar{p}, \bar{x})$. Recall that in [**17**] the generalized equation (1.1) was termed *strongly regular* at $\bar{p}$ if there exist neighborhoods U_0 of 0 in $\mathbb{R}^n$ and V_0 of 0 in $\mathbb{R}^m$ such that for the linearization

$$T(u) = D_x f(\bar{p}, \bar{x})(u) + f(\bar{p}, \bar{x}) + N(\bar{x} + u)$$

one has that $U_0 \cap T^{-1}$ is a Lipschitz continuous function on V_0. By the polyhedrality of N there is a neighborhood of the origin in the product space $\mathbb{R}^m \times \mathbb{R}^n$ such that for the mapping $F(x) = f(\bar{p}, x) + N(x)$ one has that in this neighborhood the graph of the contingent derivative of F at $(\bar{x}, 0)$ is equal to the graph of the linearization T; that is, there are neighborhoods U and V of 0 in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, such that

$$\text{gph}\, DF(\bar{x}|0) \cap (U \times V) = \text{gph}\, T \cap (U \times V). \tag{6.3}$$

The first consequence of this coincidence of the linearization and the contingent derivative is a demonstration of the equivalence between the assumption of strong regularity and the single-valuedness of $DF^{-1}(0|\bar{x})$, for the case where $f(\bar{p}, \cdot)$ is continuously differentiable. (Subinvertibility in this situation is a consequence of the fact that a polyhedral multifunction that is everywhere single-valued is a Lipschitz continuous function.)

Theorem 6.1. *In the generalized equation (1.1) let $N : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a polyhedral multifunction, and let $f : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ be a continuous function such that as a function of a single variable $f(\bar{p}, \cdot)$ is continuously differentiable on a neighborhood of $\bar{x}$, where $\bar{x}$ is a solution to (1.1) at $\bar{p}$. Define the multifunction $F(x) = f(\bar{p}, x) + N(x)$. Then (1.1) is strongly regular at $\bar{p}$ if and only if $DF^{-1}(0|\bar{x})$ is everywhere single-valued. If either of these equivalent conditions holds and f has a strong partial B-derivative in p at $(\bar{p}, \bar{x})$, then the conclusions of statements (a) and (b) of Theorem 4.1 are true for the solution mapping to this generalized equation.*

Proof. Clearly, the subinvertibility of F and the single-valuedness of $DF^{-1}(0|\bar{x})$ are all that is required, in addition to the differentiability assumptions on f, to invoke Theorem

4.1. Therefore, let us concentrate on proving the equivalence. If (1.1) is strongly regular, then we know from the implicit function theorem in [**17**] that there is a neighborhood V of 0 in $\mathbb{R}^m$ and a continuous function $x : V \to \mathbb{R}^n$ such that $x(y)$ is the unique solution to

$$y \in f(\bar{p}, x(y)) + N(x(y)).$$

This implies that F is subinvertible. Furthermore, the strong regularity requires that the inverse of the linearization T^{-1} is single-valued on a neighborhood of 0 in $\mathbb{R}^n$, and by (6.3), it follows that $DF^{-1}(0|\bar{x})$ is everywhere at most a singleton, as claimed. For the other direction, we show that $DF^{-1}(0|\bar{x})$ is Lipschitz continuous on a neighborhood of the origin; then strong regularity follows from (6.3), again. Since $f(\bar{p}, \cdot)$ is differentiable, it follows that

$$u \mapsto DF(\bar{x}|0)(u) = \nabla_x f(\bar{p}, \bar{x})u + DN(\bar{x}| - f(\bar{p}, \bar{x}))(u)$$

is polyhedral. Hence $DF^{-1}(0|\bar{x})$ is also polyhedral, and a single-valued mapping that is polyhedral is a Lipschitz continuous function. $\square$

A different implicit function theory allows us to weaken the assumption of continuous differentiability of $f(\bar{p}, \cdot)$ near $\bar{x}$ to an assumption of strong B-differentiability of $f(\bar{p}, \cdot)$ at $\bar{x}$. Here, we deploy the theorem of Robinson [**21**] and follow his suggestions in analyzing (6.1) via a reformulation as a certain system of equations. Since strong regularity is generally stronger than the twin assumptions of subinvertibility and single-valuedness of the inverse of the contingent derivative of F, and since B-differentiability of f is such a weak differentiability property, we cannot conclude the equivalence of these two sets of assumptions. Just which additional properties of f would imply equivalence is an unsettled question.

Theorem 6.2. *In the generalized equation (1.1) suppose that $N : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a maximal monotone polyhedral multifunction, and that $f : \Omega \times \mathbb{R}^n \to \mathbb{R}^n$ is a Lipschitz continuous function such that as a function of a single variable $f(\bar{p}, \cdot)$ has a strong B-derivative at $\bar{x}$, where $\bar{x}$ is a solution to (1.1) at $\bar{p}$. Define the multifunction $F(x) = f(\bar{p}, x) + N(x)$. If (1.1) is strongly regular at $\bar{p}$ then F is subinvertible at $(\bar{x}, 0)$ and $DF^{-1}(0|\bar{x})$ is everywhere at most a singleton. If, furthermore, f has a strong partial B-derivative in p at $(\bar{p}, \bar{x})$, then the conclusions of statements (a) and (b) of Theorem 4.1 are true for the solution mapping to this generalized equation.*

Proof. We proceed exactly as in the first part of Theorem 6.1. The only new thing to show is that strong regularity of (1.1) at $\bar{p}$ implies that F is subinvertible at $(\bar{x}, 0)$. Consider the simply perturbed problem

(6.4) $$y \in f(\bar{p}, x) + N(x).$$

For simplicity of notation in this proof, denote $f(\bar{p}, \cdot)$ by $\bar{p}(\cdot)$. We are going to analyze the behavior of the solution to this generalized equation by reformulating it as a system of equations

$$y = \bar{p}(n(x)) + x - n(x), \tag{6.5}$$

where $n(x)$ is the solution to the generalized equation: find $u \in \mathbb{R}^n$ to satisfy

$$0 \in u - x + N(u). \tag{6.6}$$

Since N is maximal monotone, then $n(x) = (I + N)^{-1}(x)$ is a single-valued Lipschitz continuous function on all of $\mathbb{R}^n$, by Minty's theorem [**14**]. The system (6.4) is equivalent to (6.5), since if $\hat{x}$ solves (6.5), then $n(\hat{x})$ solves (6.4), and conversely, if $\bar{x}$ solves (6.4), then $\bar{x} - \bar{p}(\bar{x})$ solves (6.5).

The solution of (6.5) can be examined using the implicit function theorem [**21, Theorem 3.2**] applied to the equation

$$g(m(x), y) = 0,$$

where

$$g((a, b), y) = \bar{p}(a) + b - y$$

and the function m is the *Minty map*

$$m(x) = (n(x), x - n(x)).$$

The requirements of this theorem are that g have a partial B-derivative with respect to $v = (a, b)$ that is strong at $v_0 = (n(\hat{x}_0), \hat{x}_0 - n(\hat{x}_0))$, where $\hat{x}_0$ solves (6.5) at $y = 0$, and that the function

$$u \mapsto D_v g(v_0, 0)[m(\hat{x}_0 + u) - m(\hat{x}_0)] \tag{6.7}$$

have an inverse that is a locally Lipschitz continuous function near 0. We now verify these two conditions.

First, let us compute the partial B-derivative of g and verify that it is strong. We apply [**21, Proposition 2.2**] and obtain

$$D_v g(v_0, 0)(a, b) = D\bar{p}(n(\hat{x}_0))(a) + b,$$

which is a strong B-derivative, since $\bar{p}$ is assumed to be strongly B-differentiable at $\bar{x} = n(\hat{x}_0)$. Next we examine the inverse of the function (6.7), which we have just computed to be

$$u \mapsto D\bar{p}(\bar{x})(n(\hat{x}_0 + u) - n(\hat{x}_0)) + u - (n(\hat{x}_0 + u) - n(\hat{x}_0)). \tag{6.8}$$

Note that the contingent derivative of n is

$$Dn(x)(u) = [I + DN(n(x)|x - n(x))]^{-1}(u),$$

and thus the equality (6.2) implies that

$$Dn(x)(u) = n(x + u) - n(x).$$

The Lipschitz continuity of the inverse of (6.8) is thus equivalent, by reversing the Minty map, to the Lipschitz continuity of the inverse of the multifunction

$$u \mapsto D\bar{p}(\bar{x})(u) + DN(\bar{x}| - \bar{p}(\bar{x}))(u),$$

and hence, the Lipschitz continuity follows from the assumption of strong regularity and equation (6.3), as in Theorem 6.1.

The two conditions of the implicit function theorem being verified, we can now conclude the existence of neighborhoods U of $\hat{x}_0$ and V of 0 and a Lipschitz continuous function $\hat{x} : V \to U$ such that $\hat{x}(0) = \hat{x}_0$ and $\hat{x}(y)$ is the unique solution to the system (6.5). Then $n(\hat{x}(\cdot))$ is a continuous function on V such that $n(\hat{x}(y))$ is the unique solution to (6.4) for all $y \in V$, with $n(\hat{x}(0)) = \bar{x}$. It follows that $\bar{p} + N$ is subinvertible at $\bar{x}$ and the proof is complete. □

7. Application to Sensitivity Analysis in Mathematical Programming.

Let us consider an optimization problem in the general form studied by Robinson in [**19**]:

$$\text{minimize} \quad h(p, x) \quad \text{subject to} \quad g(p, x) \in Q^o \text{ and } x \in C, \tag{7.1}$$

where C is a polyhedral set in $\mathbb{R}^n$ and Q^o the polar of a polyhedral cone Q in $\mathbb{R}^m$. We assume throughout this section that $h(p, \cdot)$ and $g(p, \cdot)$ are once continuously differentiable in x for all p in a neighborhood Ω of a point $\bar{p}$ in a Banach space P, and that $h(\bar{p}, \cdot)$ and $g(\bar{p}, \cdot)$ are twice continuously differentiable in x on a neighborhood of $\bar{x}$.

To fit the pattern established in the previous sections, define the function $f : \Omega \times \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ by

$$f(p, x, \mu) = (\nabla_x h(p, x) + \mu^T \nabla_x g(p, x), -g(p, x))$$

and write down the first-order necessary conditions for (7.1) as

$$0 \in f(p, x, \mu) + N_{C\times Q}(x, \mu), \tag{7.2}$$

where $N_{C\times Q}$ is the normal cone operator of convex analysis. Letting $(\bar{x}, \bar{\mu})$ be a Kuhn-Tucker point for (7.2) at $p = \bar{p}$, we first examine how assumptions standard for such optimization problems imply the subinvertibility of (7.2).

Proposition 7.1. *Assume that the generalized equation (7.2), for $p = \bar{p}$, is* regular *at $\bar{x}$, i.e.*

$$0 \in \ int \ \{g(\bar{p}, \bar{x}) + \nabla_x g(\bar{p}, \bar{x})[C - \bar{x}] - Q^o\},$$

and that the second order sufficient condition *is satisfied at the pair $(\bar{x}, \bar{\mu})$, namely $\forall u \in T_C(\bar{x})$ with u nonzero and*

$$\nabla_x g(\bar{p}, \bar{x})u \in T_{Q^o}(g(\bar{p}, \bar{x})) \text{ and } \nabla_x h(\bar{p}, \bar{x})u = 0$$

one has

$$\langle u, \nabla_x^2 L(\bar{p}, \bar{x}, \bar{\mu})u\rangle > 0,$$

where T_C is the tangent cone of convex analysis and $L(p, x, \mu) = h(p, x) + \mu^T g(p, x)$ is the Lagrangian function. Then the multifunction $f(\bar{p}, \cdot, \cdot) + N_{C\times Q}(\cdot, \cdot)$ is subinvertible at $(\bar{x}, \bar{\mu})$.

Proof. Note that by Theorem 3.2 of [**19**] there exist neighborhoods V of 0 and U of $\bar{x}$ such that the multifunction

$$SP(y) = \left\{x \in U \,\middle|\, \exists\mu \in \mathbb{R} \text{ with } y \in f(\bar{p}, x, \mu) + N_{C\times Q}(x, \mu)\right\}$$

is lower semicontinuous on V (here, the parameter y plays the role of the parameter "p" of [**19**]). Michael's selection theorem [**12**] now implies that there is a continuous selection $x(y) \in SP(y)$ on V. The multifunction

$$M(y) = \left\{\mu \in \mathbb{R}^m \,\middle|\, y \in f(\bar{p}, x(y), \mu) + N_{C\times Q}(x(y), \mu)\right\}$$

is closed by Theorem 2.3 of [**19**], is locally bounded (by regularity), and is nonempty and convex-valued. The product mapping $y \mapsto G(y) = \{x(y)\} \times M(y)$ thus fulfills the requirement of the definition of subinvertibility. □

Next, we establish B-differentiability properties of the solutions at $\bar{p}$. It is convenient to consolidate further the notation of (7.2): let $z = (x, \mu)$, let $R = C \times Q$, define $F(z) =$

$f(\bar{p}, z) + N_R(z)$, and let $\bar{z} = (\bar{x}, \bar{\mu})$ be the Kuhn-Tucker point at $p = \bar{p}$. By Remark 4.3, we may compute the contingent derivative of F as

$$DF(\bar{z}|0)(s) = D_z f(\bar{p}, \bar{z})(s) + D(N_R)(\bar{z}| - f(\bar{p}, \bar{z}))(s).$$

Applying the formulas in Rockafellar [**26**] [**27**], or King [**6**], one finds that the contingent derivative of the normal cone mapping N_R at $(\bar{z}, -f(\bar{p}, \bar{z}))$ is the normal cone operator to the *critical cone*

$$R'(\bar{z}|f(\bar{p}, \bar{z})) = \{s \in T_R(\bar{z}) \,\big|\, s \cdot f(\bar{p}, \bar{z}) = 0\}.$$

Thus $DF^{-1}(0|\bar{z})(c)$ is the solution set to the generalized equation

$$c \in D_z f(\bar{p}, \bar{z})(s) + N_{R'(\bar{z}|f(\bar{p}, \bar{z}))}(s).$$

Decoding this generalized equation yields objects that are very familiar to students of sensitivity analysis. We obtain a pair of second-order generalized equations

$$a \in \nabla_x^2 L(\bar{p}, \bar{x}, \bar{\mu})u + v^T \nabla_x g(\bar{p}, \bar{x}) + N_{C'(\bar{x}|\nabla_x L(\bar{p}, \bar{x}, \bar{\mu}))}(u) \tag{7.4}$$

$$b \in -\nabla_x g(\bar{p}, \bar{x})u + N_{Q'(\bar{\mu}|g(\bar{p}, \bar{x}))}(v), \tag{7.5}$$

which may be interpreted as the first-order necessary conditions to a certain well-known convex quadratic programming problem. According to the results of Section 6, the solutions to this pair, if unique for all left-hand sides (a, b), yield the B-derivatives of the solutions to (7.1) as p varies near $\bar{p}$. The sensitivity analysis literature ([**4**] [**11**] [**16–21**] [**29**]) describes many sorts of assumptions (linear independence, strong second-order sufficiency, etc.) that yield unique solutions for (7.4) and (7.5). We shall not discuss particular instances here—the main issue, as we have seen, is the existence of unique solutions to this pair.

Theorem 7.2. *In addition to the assumptions of Proposition 7.1, assume that the pair of generalized equations (7.4) and (7.5) have unique solutions for all left-hand sides (a, b) and that $\nabla_x h$, g, and $\nabla_x g$ are continuous functions on $\Omega \times \mathbb{R}^n$ that have strong partial B-derivatives in p at $(\bar{p}, \bar{x})$. Then there exists a neighborhood Ω of $\bar{p}$ in P and neighborhoods U and V of $\bar{x}$ and $\bar{\mu}$, respectively, such that for all $p \in \Omega$ there exist Kuhn-Tucker selections $(x(p), \mu(p)) \in U \times V$ for the perturbed version of (7.1). Furthermore, these Kuhn-Tucker selections are upper Lipschitzian and B-differentiable at $\bar{p}$, with*

$$Dx(\bar{p})(w) = u(-D_p \nabla_x L(\bar{p}, \bar{x}, \bar{\mu})(w), D_p g(\bar{p}, \bar{x})(w)),$$

and

$$D\mu(\bar{p})(w) = v(-D_p\nabla_x L(\bar{p},\bar{x},\bar{\mu})(w), D_p g(\bar{p},\bar{x})(w)),$$

where $u(a,b)$ and $v(a,b)$ are solutions to the second-order equations (7.4) and (7.5) with left-hand sides a and b. □

Proof. We shall apply Theorem 6.1 to the generalized equation (7.2). As above, let $z = (x,\mu)$, $R = C \times Q$, and define $F(z) = f(\bar{p},z) + N_R(z)$. The assumptions we have made on h and g imply in particular that f is a continuous function on $\Omega \times \mathbb{R}^{n+m}$ such that $f(\bar{p},\cdot)$ is continuously differentiable on a neighborhood of $\bar{z}$, and Proposition 7.1 implies that F is subinvertible at $(0,\bar{z})$. The above discussion and the assumption of unique solutions to (7.4) and (7.5) show that $DF^{-1}(0|\bar{z})$ is everywhere single-valued. Finally, the existence of strong partial B-derivatives of f in p at $(\bar{p},\bar{z})$ allows us to apply, via Theorem 6.1, the conclusions of statements (a) and (b) of Theorem 4.3. □

References.

1. J-P. Aubin, "Lipschitz behaviour of solutions to convex minimization problems", *Mathematics of Operations Research* **9** (1984) 97–102.
2. S. Dafermos, "Sensitivity analysis in variational inequalities", *Mathematics of Operations Research* **13** (1988) 421–434.
3. J. Dupačová and R.J-B Wets, "Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems", *Annals of Statistics* (1988) 1517–1549.
4. A.V. Fiacco, *Introduction to sensitivity analysis in nonlinear programming* (Academic Press, New York, NY, 1983).
5. S. Kakutani, "A generalization of Brouwer's fixed point theorem", *Duke Mathematics Journal* **8** (1941) 457–459.
6. A.J. King, "Asymptotic behaviour of solutions in stochastic optimization: Nonsmooth analysis and the derivation of non-normal limit distributions", Ph.D. Dissertation, University of Washington (Seattle, WA, 1986).
7. A.J. King, "Generalized delta theorems for multivalued mappings and measurable selections", *Mathematics of Operations Research* **14** (1989) 720–736.
8. A.J. King and R.T. Rockafellar, "Asymptotic theory for solutions in stochastic programming and generalized *M*-estimation", unpublished manuscript, IBM T.J. Watson Research Center (Yorktown Heights, NY, 1989).
9. B. Kummer, "Generalized equations: solvability and regularity", *Mathematical Programming Study* **21** (1984) 199–212.

10. B. Kummer, "Linearly and nonlinearly perturbed optimization problems in finite dimensions", in: J. Guddat et al, eds., *Parametric optimization and related topics*, (Akademie-Verlag, Berlin, 1987).

11. J. Kyparisis, "Sensitivity analysis framework for variational inequalities", *Mathematical Programming* **38** (1987) 203–214.

12. E. Michael, "Selected selection theorems", *American Mathematical Monthly* **63** (1956) 233–238.

13. G.J. Minty, "On the maximal domain of a 'monotone' function", *Michigan Mathematics Journal* **8** (1961) 135-137.

14. G.J. Minty, "Monotone (nonlinear) operators in Hilbert space", *Duke Mathematics Journal* (1962) 341–346.

15. R.A. Poliquin, "Proto-differentiation of subgradient set-valued mappings", unpublished manuscript, University of Alberta (Edmonton, Alberta, Canada, 1988).

16. S.M. Robinson, "Generalized equations and their solutions, part 1: basic theory", *Mathematical Programming Study* **10** (1979) 128–141.

17. S.M. Robinson, "Strongly regular generalized equations", *Mathematics of Operations Research* **5** (1980) 43–62.

18. S.M. Robinson, "Some continuity properties of polyhedral multifunctions", *Mathematical Programming Study* **14** (1981) 206–214.

19. S.M. Robinson, "Generalized equations and their solutions, part II: applications to nonlinear programming",*Mathematical Programming Study* **19** (1982) 200–221.

20. S.M. Robinson, "Local structure of feasible sets in nonlinear programming, part III: stability and sensitivity", *Mathematical Programming Study* **30** (1987) 45–66.

21. S.M. Robinson, "An implicit function theorem for B-differentiable functions", Working Paper WP-88-67, International Institute for Applied Systems Analysis (Laxenburg, Austria, 1988).

22. R.T. Rockafellar, "Local boundedness of nonlinear monotone operators", *Michigan Mathematics Journal* **16** (1969) 397-407.

23. R.T. Rockafellar, *Convex analysis* (Princeton University Press, Princeton, NJ, 1970).

24. R.T. Rockafellar, "Directional differentiability of the optimal value function in a nonlinear programming problem", *Mathematical Programming Study* **21** (1984) 213–216.

25. R.T. Rockafellar, "First and second-order epi-differentiability with applications to optimization", *Transactions of the American Mathematical Society* **307** (1988) 75–108.

26. R.T. Rockafellar, "Proto-differentiability of set-valued mappings and its applications in optimization", *Annals of the Institute of H. Poincaré: Analyse Non Lineaire* (1988) (to appear).

27. R.T. Rockafellar, "Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives", *Mathematics of Operations Research* **14** (1989) 462–484.

28. A. Shapiro, "Asymptotic distribution of test statistics in the analysis of moment structures under inequality constraints", *Biometrika* **72** (1985) 133–144.

29. A. Shapiro, "Second order sensitivity analysis and asymptotic theory of parametrized nonlinear programs", *Mathematical Programming* **33** (1985) 280–299.

30. A. Shapiro, "On concepts of directional differentiability", unpublished manuscript, University of South Africa (Pretoria, South Africa, 1986).