1-1-2001

Lipschitzian multifunctions and a Lipschitzian inverse mapping theorem

A.B. Levy

Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Lipschitzian Multifunctions and a Lipschitzian Inverse Mapping Theorem

A. B. Levy,

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 26, No. 1, February 2001, pp. 105–118
Printed in U.S.A.

LIPSCHITZIAN MULTIFUNCTIONS AND A LIPSCHITZIAN INVERSE MAPPING THEOREM

A. B. LEVY

We introduce a new class of multifunctions whose graphs under certain "kernel inverting" matrices, are locally equal to the graphs of Lipschitzian (single-valued) mappings. We characterize the existence of Lipschitzian localizations of these multifunctions in terms of a natural condition on a generalized Jacobian mapping. One corollary to our main result is a Lipschitzian inverse mapping theorem for the broad class of "max hypomonotone" multifunctions. We apply our theoretical results to the sensitivity analysis of solution mappings associated with parameterized optimization problems. In particular, we obtain new characterizations of the Lipschitzian stability of stationary points and Karush-Kuhn-Tucker pairs associated with parameterized nonlinear programs.

1. Introduction. A multifunction $S: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a mapping that takes points in $\mathbb{R}^n$ to sets in $\mathbb{R}^m$. These objects appear throughout variational analysis and optimization, and understanding when multifunctions reduce to Lipschitzian (single-valued) mappings is a basic issue in these fields. For instance, to quantify the sensitivity of solutions to a parameterized optimization problem, one could determine that the multifunction giving the set of solutions for each parameter is actually a Lipschitzian mapping. Usually, it is only necessary that a multifunction have a localization that is a Lipschitzian mapping, so we focus our attention on conditions under which there is a Lipschitzian mapping $F: \mathbb{R}^n \to \mathbb{R}^m$ whose graph agrees with the graph of the multifunction S near a point $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ ($\operatorname{gph} S$ denotes the graph of S and is the set of all pairs $(x, y) \in \mathbb{R}^{n+m}$ with $y \in S(x)$). Such a mapping F is called a *Lipschitzian localization of S near* $(\bar{x}, \bar{y})$.

An object that turns out to be crucial for our study is a generalized derivative called the strict derivative. The *strict derivative* $D_*S(\bar{x}|\bar{y}): \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ of S at $\bar{x}$ for $\bar{y}$ is defined for any $w \in \mathbb{R}^n$ to be the set of limit points of difference quotients obtained from sequences approaching $(\bar{x}, \bar{y}, w)$:

$$D_*S(\bar{x}|\bar{y})(w) = \left\{ z \,\middle|\, \begin{array}{l} \exists x^\nu \to \bar{x},\, y^\nu \in S(x^\nu),\, y^\nu \to \bar{y},\, w^\nu \to w,\, \tau^\nu \downarrow 0 \text{ with} \\ (\tilde{y}^\nu - y^\nu)/\tau^\nu \to z \text{ for some } \tilde{y}^\nu \in S(x^\nu + \tau^\nu w^\nu) \end{array} \right\}$$

When the multifunction happens to be a Lipschitzian mapping $F: \mathbb{R}^n \to \mathbb{R}^m$ near $\bar{x}$, the strict derivative at $\bar{x}$ is written as $D_*F(\bar{x})$ and can be obtained by considering only sequences x^ν approaching $\bar{x}$:

$$D_*F(\bar{x})(w) = \left\{ z \mid \exists x^\nu \to \bar{x},\, \tau^\nu \downarrow 0 \text{ with } \frac{F(x^\nu + \tau^\nu w) - F(x^\nu)}{\tau^\nu} \to z \right\}$$

For Lipschitzian mappings F, single-valuedness of the strict derivative mapping $D_*F(\bar{x})$ corresponds to the classical notion of strict differentiability (hence the name "strict derivative mapping"), and for continuously differentiable mappings, the strict derivative coincides with the Jacobian. The strict derivative has been studied before (e.g. Rockafellar and Wets 1998 or Kummer 1991 where it is called "Thibault's directional derivative.)"

Received February 2, 1998; revised July 20, 2000.
AMS 2000 subject classification. Primary: 90C31; secondary: 90C30.
OR/MS subject classification. Primary: Programming/nonlinear.
Key words. Kernel inverting matrix, Lipschitzian manifold, solution mapping, strict derivative, hypomonotonicity, sensitivity analysis, variational condition, nonlinear programming.

0364-765X/01/2601/0105/$05.00
1526-5471 electronic ISSN, © 2001, INFORMS

There are already at least two different characterizations for the existence of a Lipschitzian localization of a general multifunction $S:\ {}^n \rightrightarrows\ {}^m$; namely Levy and Poliquin 1997, Theorem 2.1, and Rockafellar and Wets (1998, Theorem 9.51). The second of these utilizes the strict derivative mapping, and they both involve a generalized Lipschitzian property for multifunctions called "Aubin continuity." A multifunction $S:\ {}^n \rightrightarrows\ {}^m$ is *Aubin continuous* at $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ if there are bounded neighborhoods X of $\bar{x}$ and Y of $\bar{y}$, as well as a positive scalar L such that the inclusion

$$S(x'') \cap Y \subseteq S(x') + L \parallel x' - x'' \parallel \mathbb{B}_m \tag{1}$$

holds for all x' and x'' in X (here $\mathbb{B}_m$ denotes the unit ball in ^{m}). It is easy to see that a single-valued mapping $F:\ {}^n \to\ {}^m$ is Aubin continuous at $(\bar{x}, F(\bar{x}))$ if and only if F is Lipschitzian at $\bar{x}$. The two characterizations Levy and Poliquin (1997, Theorem 2.1) and Rockafellar and Wets (1998, Theorem 9.51) are combined in the following theorem.

THEOREM 1.1. COMBINATION OF LEVY AND POLIQUIN 1997, THEOREM 2.1, AND ROCKAFELLAR AND WETS 1998, THEOREM 9.51). *For a multifunction $S:\ {}^n \rightrightarrows\ {}^m$ and a pair $(\bar{x}, \bar{y}) \in \operatorname{gph} S$, the following are equivalent*:

(a) *S has a Lipschitzian localization near $(\bar{x}, \bar{y})$.*
(b) *S is Aubin continuous at $(\bar{x}, \bar{y})$ with $D_*S(\bar{x}|\bar{y})(0) = \{0\}$.*
(c) *S is Aubin continuous at $(\bar{x}, \bar{y})$ and max hypomonotone near $(\bar{x}, \bar{y})$.*

The monotonicity property in (c) involves the existence of a maximal monotone linear perturbation of S. Speci cally, a multifunction $S:\ {}^n \rightrightarrows\ {}^n$ is *max hypomonotone near* $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ if there exists a nonnegative multiple r of the identity mapping $I:\ {}^n \to\ {}^n$ such that the multifunction $S + rI$ is maximal monotone in the usual sense near $(\bar{x}, \bar{y} + r\bar{x})$. Max hypomonotonicity for multifunctions from one Euclidean space to a different one is de ned in Levy and Poliquin (1997) by trivially extending the multifunction so that it maps between the same spaces. Max hypomonotonicity is weaker than maximal monotonicity (which is just max hypomonotonicity with $r = 0$) and many multifunctions exhibit this property. All Lipschitzian mappings are max hypomonotone (cf. Levy and Poliquin 1997 and the fact that any monotone single-valued mapping is maximal monotone), and many optimality conditions can be interpreted as nding the zeroes of max hypomonotone multifunctions. (See Levy and Poliquin 1997 for more on max hypomonotone multifunctions.)

Our aim in the present paper is to show that for a very large class of multifunctions, Aubin continuity is automatic under the "kernel condition" in (b) on the strict derivative. The class of multifunctions on which we focus is previously unidenti ed and consists of those multifunctions whose graphs are "kernel inverting Lipschitzian manifolds." This de nition involves the *kernel inverting* matrices $A \in\ {}^{2n\times 2n}$ which are distinguished by having invertible $n \times n$ submatrices in the upper right and lower left quadrants, and a zero $n \times n$ submatrix in the upper left quadrant. The graph of a multifunction $S:\ {}^n \rightrightarrows\ {}^n$ is a *kernel inverting Lipschitzian manifold near* $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ if there exists a neighborhood U of $(\bar{x}, \bar{y})$ in the graph space ^{2n}, and a Lipschitzian mapping $F:\ {}^n \to\ {}^n$ for which $\operatorname{gph} S \cap U = A \operatorname{gph} F) \cap U$ for some kernel inverting matrix $A \in\ {}^{2n\times 2n}$. This class includes all Lipschitzian mappings, all max hypomonotone multifunctions, and even all multifunctions whose inverses are max hypomonotone. Our main theorem has the following statement.

THEOREM 1.2. *For a multifunction $S:\ {}^n \rightrightarrows\ {}^n$ and a pair $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ near which $\operatorname{gph} S$ is a kernel inverting Lipschitzian manifold, the following are equivalent*:

(a) *S has a Lipschitzian localization near $(\bar{x}, \bar{y})$.*
(b) $D_*S(\bar{x}|\bar{y})(0) = \{0\}$.

This characterization is in some sense the best we can do, as it is not possible to drop the kernel condition (b) above in favor of the Aubin continuity property. For example, consider the mapping $S: \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ whose inverse mapping S^{-1} assigns to a point with polar coordinates (r, θ) the point $(r, 2\theta)$. Clearly, $S(0) = 0$ but S is double-valued for all points near 0. In spite of this double-valuedness, S is Aubin continuous at $(0, 0)$ (see Rockafeller and Wets 1998). As dictated by Theorem 1.2 however, the strict derivative set $D_*S(0, 0)(0)$ does not reduce to the point 0, but in fact equals the whole space $\mathbb{R}^2$.

Notice also that our characterization in Theorem 1.2 only includes multifunctions S from one Euclidean space to itself whereas the characterizations recorded in Theorem 1.1 extend to multifunctions between any Euclidean spaces. Our focus on multifunctions between the same Euclidean spaces is necessary to obtain our results, and moreover does not necessarily restrict our application of this result to multifunctions between the same spaces (as we show in §4).

In §3, we study max hypomonotonicity and show that multifunctions S that are max hypomonotone themselves, as well as those multifunctions having max hypomonotone inverses S^{-1}, are both examples of multifunctions for which $\operatorname{gph} S$ is a kernel inverting Lipschitzian manifold. This fact allows us to prove the following Lipschitzian inverse mapping theorem for max hypomonotone multifunctions as a corollary to Theorem 1.2.

THEOREM 1.3. *For a multifunction $S: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ that is max hypomonotone near $(\bar{x}, \bar{y}) \in \operatorname{gph} S$, the following are equivalent*:

(a) *S^{-1} has a Lipschitzian localization near $(\bar{y}, \bar{x})$.*

(b) *$0 \in D_*S(\bar{x}|\bar{y})(x)$ only for $x = 0$.*

Since any Lipschitzian mapping is necessarily max hypomonotone, our inverse mapping theorem is a direct extension of Kummer's Lipschitzian inverse mapping theorem for Lipschitzian mappings (Kummer 1991, Theorem 1.1). Kummer's (1991) Theorem 1.1 says that the inverse of a Lipschitzian mapping $F: \mathbb{R}^n \to \mathbb{R}^n$ has a Lipschitzian localization if and only if $0 \in D_*F(\bar{x})(x)$ only for $x = 0$. This result (and consequently our own Theorem 1.3) is a natural generalization of the classical inverse mapping theorem, since the strict derivative reduces to the Jacobian when F is differentiable. However, our Lipschitzian inverse mapping theorem not only extends the classical inverse mapping theorem, but extends it very broadly, and with many useful consequences that do not follow from Kummer (1991, Theorem 1.1).

In the final two sections of this paper, we explore some of these consequences as we apply Theorem 1.3 to analyze the sensitivity of solutions to parameterized nonlinear programs. We obtain characterizations of the Lipschitzian stability of stationary points as well as Karush-Kuhn-Tucker pairs associated with the nonlinear programs. The result about the stationary points is unprecedented, and the result about the KKT pairs is different from any of the previously available characterizations. Moreover, our method for obtaining the latter characterization is new and straightforward, involving only the computation of a strict derivative. Since the theory developed here is quite general, many other applications to sensitivity analysis in optimization could be made in the same manner, though we will not pursue these in the present paper.

2. Kernel inverting Lipschitzian manifolds.

The kernel inverting matrices $A \in \mathbb{R}^{2n\times 2n}$ have the form

$$A = \begin{bmatrix} 0 & A_2 \\ A_3 & A_4 \end{bmatrix},$$

for invertible $n \times n$ matrices A_2 and A_3. According to this structure, kernel inverting matrices are always invertible with inverse given by

$$A^{-1} = \begin{bmatrix} (-A_3^{-1}A_4A_2^{-1}) & A_3^{-1} \\ A_2^{-1} & 0 \end{bmatrix}$$

In fact, it is easy to see that the kernel inverting matrices are the same as the class of invertible matrices $A \in \mathbb{R}^{2n\times 2n}$ with zero principal $n \times n$ submatrix; the invertibility of the $n \times n$ submatrices A_2 and A_3 follows from the invertibility of A.

Recall that gph S is a kernel inverting Lipschitzian manifold near $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ if there exists a Lipschitzian mapping $F: \mathbb{R}^n \to \mathbb{R}^n$ such that nearby $(\bar{x}, \bar{y})$ in the graph space, gph S and A gph F are the same. If this is the situation, and owing to the invertibility of A, gph S is also an "n-dimensional Lipschitzian manifold near $(\bar{x}, \bar{y})$" in the sense of Rockafellar (1985). (See Levy and Rockafellar 1996 for a discussion of these objects.) General n-dimensional Lipschitzian manifolds represent all the local transformations, smooth in both directions, of the graphs of Lipschitzian mappings. Kernel inverting Lipschitzian manifolds are then special cases of n-dimensional Lipschitzian manifolds, where the local transformations are linear and have the special form A described above. If gph S is an n-dimensional Lipschitzian manifold, then gph S^{-1} is too since these graph sets are the same under the invertible linear transformation

$$\tilde{I} := \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$$

Moreover, the same Lipschitzian mapping can be associated with both S and S^{-1}. This is not true for kernel inverting Lipschitzian manifolds, as the matrix $\tilde{I}$ does not necessarily transform one kernel inverting matrix into another. Nevertheless, we show in Theorem 3.1 that the sets gph S and gph S^{-1} are both kernel inverting Lipschitzian manifolds for max hypomonotone multifunctions S.

The special form of the linear transformation A provides the motivation for the label "kernel inverting," as will be clear after the following lemma is established.

LEMMA 2.1. *Let $S: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a multifunction having* gph S *a kernel inverting Lipschitzian manifold near $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ (so* gph S *is equal to A* gph F *near $(\bar{x}, \bar{y})$), and let $\bar{u} \in \mathbb{R}^n$ be the base point in the domain of F corresponding to $(\bar{x}, \bar{y})$ (so $(\bar{u}, F(\bar{u})) = A^{-1}(\bar{x}, \bar{y})$). Then the graph of the strict derivative mapping $D_*S(\bar{x}|\bar{y})$ satisfies*

$$\operatorname{gph} D_*S(\bar{x}|\bar{y}) = A \operatorname{gph} D_*F(\bar{u}) \tag{2}$$

Moreover, the vectors in the kernel of the strict derivative $D_*F(\bar{u})$ are the same as the vectors obtained by applying the matrix A_3^{-1} (the inverse of the $n \times n$ matrix in the lower left quadrant of the matrix A) to the vectors in the kernel of the strict derivative of the inverse of S:

$$0 \in D_*F(\bar{u})(u) \Longleftrightarrow A_3 u \in D_*S(\bar{x}|\bar{y})(0) \Longleftrightarrow 0 \in D_*S^{-1}(\bar{y}|\bar{x})(A_3 u)$$

PROOF. According to the definition of the strict derivative mapping, its graph gph $D_*S(\bar{x}|\bar{y})$ is the outer limit (in the sense of Painlevé-Kuratowski, see for example Levy and Rockafellar 1995) of the sets

$$\frac{\operatorname{gph} S - (x, y)}{t}$$

for sequences of positive scalars $t \downarrow 0$ and pairs $(x, y) \in \operatorname{gph} S$ with $(x, y) \to (\bar{x}, \bar{y})$. Since gph S equals A gph F near $(\bar{x}, \bar{y})$, this outer limit is the same as the outer limit of the sets $(A \operatorname{gph} F - (x, y))/t$. Since the matrix A is invertible, this set limit is the same as the outer limit of the sets

$$\frac{A\,(\operatorname{gph} F - (u, F(u))}{t}$$

for sequences of points $u \in \mathbb{R}^n$ converging to $\bar{u}$. Finally, according to the definition of the strict derivative of F, this limit set is just A gph $D_*F(\bar{u})$.

To prove our assertion about the kernels of the strict derivative mappings, we appeal to identity (2) and see that $u \in \mathbb{R}^n$ satisfies $(u, 0) \in \operatorname{gph} D_*F(\bar{u})$ if and only if $A(u, 0) \in$

gph $D_*S(\bar{x}|\bar{y})$. The form of the kernel inverting matrix A ensures that this occurs if and only if $(0, A_3u) \in \operatorname{gph} D_*S(\bar{x}|\bar{y})$, which is the meaning of the first equivalence above. The second equivalence follows directly from the definition of the inverse mapping (see for example Rockafellar and Wets 1998, Chapter 9).

Lemma 2.1 shows that strict derivatives preserve the graphical relationship inherent in multifunctions whose graphs are kernel inverting Lipschitzian manifolds. (In fact, the first part of the proof of Lemma 2.1 works for multifunctions whose graphs are any kind of Lipschitzian manifold.) The final statement in Lemma 2.1 provides the motivation for the term "kernel inverting": Even though the transformation A maps gph F to gph S and gph $D_*F(\bar{u})$ to gph $D_*S(\bar{x}|\bar{y})$, it "inverts" the kernel of $D_*F(\bar{u})$ by mapping it via A_3 to the kernel of the strict derivative of the *inverse* of S.

The next lemma shows that multifunctions whose graphs are kernel inverting Lipschitzian manifolds, can actually be expressed directly in terms of the inverse multifunctions obtained from their associated Lipschitzian mapping.

LEMMA 2.2. *Let* $S: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *be a multifunction whose graph is a kernel inverting Lipschitzian manifold near* $(\bar{x}, \bar{y}) \in \operatorname{gph} S$:

$$U \cap \operatorname{gph} S = U \cap (A \operatorname{gph} F) = U \cap \begin{bmatrix} 0 & A_2 \\ A_3 & A_4 \end{bmatrix} \operatorname{gph} F,$$

and let the mapping $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *be defined by* $G = (A_4 + A_3 F^{-1}) A_2^{-1}$. *Then the set* $\operatorname{gph} S \cap U$ *agrees with the set* $\operatorname{gph} G \cap U$, *and* S *has a Lipschitzian localization near* $(\bar{x}, \bar{y})$ *if and only if* F^{-1} *has a Lipschitzian localization near* $(F(\bar{u}), \bar{u})$ (*where* $\bar{u} \in \mathbb{R}^n$ *is the base point in the domain of* F *corresponding to* $(\bar{x}, \bar{y})$).

PROOF. The local identity between the graphs of S and G is immediate since the set $A \operatorname{gph} F$ is identical to the set

$$\begin{bmatrix} A_2 & 0 \\ A_4 & A_3 \end{bmatrix} \operatorname{gph} F^{-1}$$

The claim about the Lipschitzian localizations follows because A_3 is invertible.

According to Lemma 2.2, there is a neighborhood $Y \subseteq \mathbb{R}^n$ of $\bar{y}$ such that the images $S(x) \cap Y$ are the same as the images $G(x) \cap Y$ for all $x \in \mathbb{R}^n$ near $\bar{x}$. Therefore according to the definition of G, much about the local behavior of the multifunction S is dictated by the local behavior of the inverse of the Lipschitzian mapping F. This fact together with Kummer (1991, Theorem 1.1) allows us to prove our Theorem 1.2.

PROOF OF THEOREM 1.2. According to Lemma 2.2, the existence of a Lipschitzian localization for S near $(\bar{x}, \bar{y})$ is equivalent to the existence of a Lipschitzian localization for F^{-1} near $(F(\bar{u}), \bar{u})$ (where $(\bar{u}, F(\bar{u})) = A^{-1}(\bar{x}, \bar{y})$). By Kummer's inverse Lipschitzian mapping theorem (Kummer 1991, Theorem 1.1), this is equivalent to the condition that $0 \in D_*F(\bar{u})(u)$ only for $u = 0$. Since the lower left $n \times n$ submatrix A_3 is invertible, Lemma 2.1 gives us that this final condition is equivalent to $D_*S(\bar{x}|\bar{y})(0) = \{0\}$.

The proof of Theorem 1.2 ultimately depends only on Kummer (1991, Theorem 1.1), once the relationship between S and the mapping F^{-1} is established in Lemma 2.2. This highlights the importance of the relationship between S and F^{-1}, and establishes this relationship as the essential property of multifunctions whose graphs are kernel inverting Lipschitzian manifolds. For any multifunction whose graph is a Lipschitzian manifold, there is an associated Lipschitzian mapping, but this association is only between the graphs of these mappings. However, multifunctions whose graphs are kernel inverting Lipschitzian manifolds can actually be expressed directly in terms of the inverses of their associated Lipschitzian mappings via Lemma 2.2. This additional property has powerful consequences for these multifunctions, as their behavior can be very closely linked to that of their associated Lipschitzian mappings. These results would not be very important however if there was a scarcity of multifunctions whose graphs are kernel inverting Lipschitzian manifolds. The exciting discovery of this paper is not only that these multifunctions have useful structure like that exploited to obtain Theorem 1.2, but that these multifunctions are very common in optimization and variational analysis. This latter topic is the focus of the remaining two sections of this paper.

3. Max hypomonotone multifunctions. In the Introduction, we discussed multifunctions which possess maximal monotone linear perturbations. In this section we show that the class of multifunctions whose graphs are kernel inverting Lipschitzian manifolds includes any multifunction that has a maximal monotone linear perturbation or whose inverse has such a perturbation. We recall that a multifunction $S: \ {}^n \rightrightarrows \ {}^n$ is said to be *max hypomonotone near* $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ if there exists a nonnegative multiple r of the identity mapping $I: \ {}^n \rightrightarrows \ {}^n$ such that the multifunction $S + rI$ is maximal monotone near $(\bar{x}, \bar{y} + r\bar{x})$. The monotonicity part of this property requires that the inequality

$$\langle x - x', (y + rx) - (y' + rx') \rangle \geq 0,$$

be satis ed for all x and x' near $\bar{x}$, and all y and y' near $\bar{y}$ satisfying $y \in S(x)$ and $y' \in S(x')$. The maximality demands that near $(\bar{x}, \bar{y} + r\bar{x})$, no other monotone multifunction contains the graph of $S + rI$ in its graph. If $r = 0$, this is just the usual maximal monotonicity property near $(\bar{x}, \bar{y})$. All Lipschitzian mappings are max hypomonotone (cf. Levy and Poliquin 1997 and the fact that any monotone single-valued mapping is maximal monotone), as well as many common multifunctions in variational analysis.

The notion of max hypomonotonicity has already received some attention for subdifferential multifunctions $\partial f: \ {}^n \rightrightarrows \ {}^n$, where $\partial f(x)$ denotes the set of limiting proximal subgradients of the function $f: \ {}^n \to \ \cup \{\infty\}$ at x. Some of this work has focussed on multifunctions ∂f associated with "subdifferentially continuous" functions: f is *subdifferentially continuous at* $\bar{x}$ *for* $\bar{v} \in \partial f(\bar{x})$ if $f(x)$ is close to $f(\bar{x})$ whenever x is close to $\bar{x}$ and there exists an element $v \in \partial f(x)$ close to $\bar{v}$ (see Poliquin and Rockafellar 1995). The class of subdifferentially continuous functions covers many of the functions in optimization including all convex functions, all lower-$\mathscr{C}^2$ functions, and all primal lower-nice functions (see Poliquin 1991 for a discussion of primal lower-nice functions). In particular, "strongly amenable" functions (compositions of convex functions with C^2 mappings under a constraint quali cation, see Poliquin and Rockafellar 1995) are primal lower-nice and hence subdifferentially continuous. Poliquin and Rockafellar (1995) introduced the class of "prox-regular" functions, and showed that this class also includes all convex functions, all lower-$\mathscr{C}^2$ functions, and all primal lower-nice functions. A subdifferentially continuous function f is called *prox-regular at* $\bar{x}$ *for* $\bar{y} \in \partial f(\bar{x})$ if ∂f is max hypomonotone near $(\bar{x}, \bar{y})$. Thus many of the most common multifunctions in optimization (i.e., the subdifferentials of prox-regular functions) are locally max hypomonotone. One particular example of such a multifunction that appears in parameterized optimization is the normal cone mapping N_C associated with a set $C \subseteq \ {}^n$ de ned by the smooth functions g_i as follows:

$$C := \{x : g_1(x) \leq 0, \ldots, g_s(x) \leq 0, g_{s+1}(x) = 0, \ldots, g_m(x) = 0\}$$

The normal cone mapping N_C is just the subdifferential of the indicator function δ_C associated with C (the indicator function is zero for points in C and is empty-valued for points not in C). Under the usual Mangasarian-Fromovitz constraint quali cation, the indicator function δ_C is prox-regular and subdifferentially continuous (see Poliquin and Rockafellar 1995), so its subdifferential mapping N_C is max hypomonotone.

The following lemma provides an association between locally max hypomonotone multifunctions and Lipschitzian mappings. This association will be the basis for showing that locally max hypomonotone multifunctions have graphs that are kernel inverting Lipschitzian manifolds.

Lemma 3.1. *If a multifunction* $S:\ ^n \rightrightarrows\ ^n$ *is max hypomonotone near* $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ *(with* $r \geq 0$*), then for any* $c > r$*, the mapping* $F := (S + cI)^{-1}$ *has a Lipschitzian localization near* $\bar{y} + c\bar{x}$*.*

Proof. According to the hypomonotonicity of S and the de nition of the mapping F, any elements $(y, x) \in \operatorname{gph} F$ and $(y', x') \in \operatorname{gph} F$ close enough to $(\bar{y} + c\bar{x}, \bar{x})$, satisfy the inequality

$$\langle y - y', x - x' \rangle \geq (c - r) \parallel x - x' \parallel^2$$

This in turn leads to the inequality

$$\parallel x - x' \parallel \leq \frac{1}{c - r} \parallel y - y' \parallel, \tag{3}$$

which gives that $x = F(y)$ and $x' = F(y')$. The maximality of the hypomonotonicity of S ensures that F is actually a single-valued mapping with values for all y near $\bar{y} + c\bar{x}$. Finally, from the bound in (3), we conclude that F is Lipschitzian near $\bar{y} + c\bar{x}$ as claimed.

Another result which will be useful is the following lemma due partly to Poliquin and Rockafellar (1995).

Lemma 3.2 Poliquin and Rockafellar 1995, Lemma 4.5). *For any multifunction* $S:\ ^n \rightrightarrows\ ^n$ *and any* $c > 0$*, one has the identity*

$$c^{-1}\left[I - (I + cS^{-1})^{-1}\right] = (S + cI)^{-1}$$

Moreover, the mapping $(I + cS^{-1})^{-1}$ *has a Lipschitzian localization near* $\bar{y} + c\bar{x}$ *if and only if the mapping* $(S + cI)^{-1}$ *has a Lipschitzian localization near* $\bar{y} + c\bar{x}$*.*

Proof. The identity comes from Poliquin and Rockafellar (1995, Lemma 4.5), and the observation about the Lipschitzian localizations follows from the identity.

Having furnished Lemmas 3.1 and 3.2, we can now prove that for any multifunction $S:\ ^n \rightrightarrows\ ^n$ which is max hypomonotone or has max hypomonotone inverse, $\operatorname{gph} S$ is a kernel inverting Lipschitzian manifold.

Theorem 3.1. *For a multifunction* $S:\ ^n \rightrightarrows\ ^n$ *that is max hypomonotone near* $(\bar{x}, \bar{y}) \in \operatorname{gph} S$*, the set* $\operatorname{gph} S$ *is a kernel inverting Lipschitzian manifold near* $(\bar{x}, \bar{y})$*, and the set* $\operatorname{gph} S^{-1}$ *is a kernel inverting Lipschitzian manifold near* $(\bar{y}, \bar{x})$*.*

Proof. We rst show that the max hypomonotonicity of S implies that $\operatorname{gph} S$ is a kernel inverting Lipschitzian manifold near $(\bar{x}, \bar{y})$. According to Lemma 3.1, there is a Lipschitzian mapping F that agrees with the mapping $(S + cI)^{-1}$ near $\bar{y} + c\bar{x}$. However, the graph of S is equal to $A \operatorname{gph}(S + cI)^{-1}$ for the kernel inverting matrix A given by

$$A = \begin{bmatrix} 0 & I \\ I & -cI \end{bmatrix}$$

Therefore, $\operatorname{gph} S$ is a kernel inverting Lipschitzian manifold near $(\bar{x}, \bar{y})$.

Now we show that $\operatorname{gph} S^{-1}$ is a kernel inverting Lipschitzian manifold near $(\bar{y}, \bar{x})$. According to Lemmas 3.1 and 3.2, there is a Lipschitzian mapping $F:\ ^n \rightarrow\ ^n$ that agrees with $(I + cS^{-1})^{-1}$ near $\bar{y} + c\bar{x}$. Because the graph of S^{-1} is the same as the graph of $(I + cS^{-1})^{-1}$ under the matrix

$$A = \begin{bmatrix} 0 & I \\ I/c & -I/c \end{bmatrix},$$

this translates into the existence of a neighborhood $U \subseteq\ ^{2n}$ of $(\bar{y}, \bar{x})$ for which the identity $\operatorname{gph} S^{-1} \cap U = (A \operatorname{gph} F) \cap U$ holds. The matrix A is clearly a kernel inverting matrix, so the graph of the inverse multifunction S^{-1} is a kernel inverting Lipschitzian manifold near $(\bar{y}, \bar{x})$.

As we mentioned in the previous section, the result in Theorem 3.1 does not follow because the sets gph S and gph S^{-1} are the same under the invertible linear transformation $\tilde{I}$. Moreover, the Lipschitzian mappings associated with S and S^{-1} are different, but they are related according to the formula in Lemma 3.2.

Since the subdifferentials of prox-regular functions are max hypomonotone, we get the following corollary immediately from Theorem 3.1.

COROLLARY 3.1. *For any subdifferentially continuous function* $f : {}^n \to \ \cup \{\infty\}$ *that is prox-regular at* $\bar{x}$ *for* $\bar{y} \in \partial f(\bar{x})$, *the set* gph ∂f *is a kernel inverting Lipschitzian manifold near* $(\bar{x}, \bar{y})$, *and the set* gph ∂f^{-1} *is a kernel inverting Lipschitzian manifold near* $(\bar{y}, \bar{x})$.

We now use our Theorem 1.2 to prove our inverse mapping theorem for max hypomonotone multifunctions.

PROOF OF THEOREM 1.3. This result follows immediately from Theorem 1.2 applied to the inverse multifunction S^{-1}, since the graph of the inverse multifunction S^{-1} is a kernel inverting Lipschitzian manifold near $(\bar{y}, \bar{x})$ according to Theorem 3.1. The kernel condition in (b) of Theorem 1.3 is equivalent to the condition that $D_* S^{-1}(\bar{y}|\bar{x})(0) = \{0\}$.

As we mentioned in the Introduction, Theorem 1.3 is a direct extension to max hypomonotone multifunctions of the classical inverse mapping theorem. Again, since the subdifferentials of prox-regular functions are max hypomonotone, we get the following immediate corollary to Theorem 1.3.

COROLLARY 3.2. *For any subdifferentially continuous function* $f : {}^n \to \ \cup \{\infty\}$ *that is prox-regular at* $\bar{x}$ *for* $\bar{y} \in \partial f(\bar{x})$, *the following are equivalent*:

(a) ∂f^{-1} *has a Lipschitzian localization near* $(\bar{y}, \bar{x})$.

(b) $0 \in D_* \ \partial f)(\bar{x}|\bar{y})(x)$ *only for* $x = 0$.

The following is the obvious complement to Theorem 1.3.

THEOREM 3.2. *For a multifunction* $S : {}^n \rightrightarrows {}^n$ *that is max hypomonotone near* $(\bar{x}, \bar{y}) \in$ gph S, *the following are equivalent*:

(a) S *has a Lipschitzian localization near* $(\bar{x}, \bar{y})$.

(b) $D_* S(\bar{x}|\bar{y})(0) = \{0\}$.

(c) S *is Aubin continuous at* $(\bar{x}, \bar{y})$.

PROOF. Conditions (a) and (c) are equivalent according to Theorem 1.1. The equivalence of (a) and (b) follows from Theorem 1.2 since gph S is a kernel inverting Lipschitzian manifold near $(\bar{x}, \bar{y})$ according to Theorem 3.1.

Notice that in the setting of Theorem 3.2, the kernel condition on the strict derivative (b) and the Aubin continuity condition (c) are equivalent, which is not the case in the more general situation covered in Theorem 1.2. Recall the example from the Introduction where the mapping $S : {}^2 \rightrightarrows {}^2$ was de ned by its inverse mapping S^{-1} which assigns to a point with polar coordinates (r, θ), the point $(r, 2\theta)$. Clearly, $S(0) = 0$ but S is double-valued for all points near 0. In spite of this double-valuedness, S is Aubin continuous near $(0, 0)$. In this case, S^{-1} is max hypomonotone near $(0, 0)$, but the multifunction S is not max hypomonotone near $(0, 0)$. Therefore, the characterization (a) $\Leftrightarrow$ (c) in Theorem 3.2 does not apply in this case, even though the characterization (a) $\Leftrightarrow$ (b) does apply here (cf. Theorem 1.2).

A nal consequence of our Theorem 1.2 when combined with Theorem 1.1, is the following suf cient condition for max hypomonotonicity in certain multifunctions.

COROLLARY 3.3. *For a multifunction* $S : {}^n \rightrightarrows {}^n$ *and a pair* $(\bar{x}, \bar{y}) \in$ gph S, *if either* (i) gph S *is a kernel inverting Lipschitzian manifold near* $(\bar{x}, \bar{y})$, *or if* (ii) S *is Aubin continuous at* $(\bar{x}, \bar{y})$, *then* S *is max hypomonotone near* $(\bar{x}, \bar{y})$ *as long as* $D_* S(\bar{x}|\bar{y})(0) = \{0\}$.

4. Applications to solution multifunctions. One type of multifunction whose local Lipschitzian properties we would like to characterize, is a "solution multifunction" S defined by

$$S(z, p) := \{x : z \in G(x, p) + N(x)\}, \tag{4}$$

for some multifunction $N : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ that is max hypomonotone near $\bar{x}$, and a mapping $G : \mathbb{R}^{n+d} \to \mathbb{R}^n$ which is Lipschitzian in p uniformly in x around $(\bar{x}, \bar{p})$ and which is differentiable with respect to x, with partial Jacobian mapping $\nabla_x G(x, p)$ depending continuously on (x, p) in a neighborhood of $(\bar{x}, \bar{p})$. The condition $z \in G(x, p) + N(x)$ is a type of generalized variational inequality called a *variational condition*. Many kinds of optimality conditions can be formulated in terms of variational conditions, so solution multifunctions of the type (4) can represent solutions to these optimality conditions (see Levy and Rockaffellar 1995 for more on variational conditions).

Levy and Poliquin (1997, Theorem 4.1) characterized the existence of Lipschitzian localizations for these solution multifunctions in terms of Lipschitzian localizations of the "linearization multifunction" $L_S : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined by

$$L_S(z) := \{x : z \in G(\bar{x}, \bar{p}) + \nabla_x G(\bar{x}, \bar{p})(x - \bar{x}) + N(x)\} \tag{5}$$

It is easy to see that the inverse of the linearization multifunction is just the multifunction $x \mapsto G(\bar{x}, \bar{p}) + \nabla_x G(\bar{x}, \bar{p})(x - \bar{x}) + N(x)$, which is max hypomonotone near $\bar{x}$ since N is. Another useful property of the linearization multifunction is that its strict derivative is always contained in the strict derivative of S.

Lemma 4.1. *For $\bar{x}$ and $\bar{p}$ as above, and for $\bar{z} \in \mathbb{R}^n$ with $\bar{x} \in L_S(\bar{z})$, the following relationships hold for any $w \in \mathbb{R}^n$:*

$$\Big\{v : w - \nabla_x G(\bar{x}, \bar{p})(v) \in D_*N\big(\bar{x}|\bar{z} - G(\bar{x}, \bar{p})\big)(v)\Big\}$$
$$= D_*L_S(\bar{z}|\bar{x})(w) \subseteq D_*S(\bar{z}, \bar{p}|\bar{x})(w)$$

Proof. First we show the inclusion

$$\Big\{v : w - \nabla_x G(\bar{x}, \bar{p})(v) \in D_*N\big(\bar{x}|\bar{z} - G(\bar{x}, \bar{p})\big)(v)\Big\} \subseteq D_*L_S(\bar{z}|\bar{x})(w) \tag{6}$$

For any pair (w, v) with $w - \nabla_x G(\bar{x}, \bar{p})(v)$ in the set $D_*N\big(\bar{x}|\bar{z} - G(\bar{x}, \bar{p})\big)(v)$, there are sequences $x^\nu \to \bar{x}$, $z^\nu \to \bar{z} - G(\bar{x}, \bar{p})$, $v^\nu \to v$, and $t^\nu \downarrow 0$ with $z^\nu \in N(x^\nu)$ and $\tilde{z}^\nu \in N(x^\nu + t^\nu v^\nu)$ satisfying

$$w^\nu := \frac{\tilde{z}^\nu - z^\nu}{t^\nu} \to w - \nabla_x G(\bar{x}, \bar{p})(v)$$

It follows that x^ν satisfies the two inclusions

$$x^\nu \in L_S\Big(z^\nu + G(\bar{x}, \bar{p}) + \nabla_x G(\bar{x}, \bar{p})(x^\nu - \bar{x})\Big),$$
$$x^\nu + t^\nu v^\nu \in L_S\Big(z^\nu + G(\bar{x}, \bar{p}) + \nabla_x G(\bar{x}, \bar{p})(x^\nu - \bar{x}) + t^\nu\big(w^\nu + \nabla_x G(\bar{x}, \bar{p})(v^\nu)\big)\Big)$$

From these two inclusions, it follows that v is an element of the set $D_*L_S(\bar{z}|\bar{x})(w)$, and the inclusion (6) is established.

Now we consider a generic element $v \in D_*L_S(\bar{z}|\bar{x})(w)$, which entails the existence of sequences $z^\nu \to \bar{z}$, $x^\nu \to \bar{x}$, $w^\nu \to w$, and $t^\nu \downarrow 0$ with $x^\nu \in L_S(z^\nu)$ and $\tilde{x}^\nu \in L_S(z^\nu + t^\nu w^\nu)$ satisfying $(\tilde{x}^\nu - x^\nu)/t^\nu \to v$. Let ϵ^ν be defined by

$$\epsilon^\nu := \frac{G(x^\nu, \bar{p}) - G(\tilde{x}^\nu, \bar{p}) - \nabla_x G(\bar{x}, \bar{p})(x^\nu - \tilde{x}^\nu)}{t^\nu}$$

Since the quotient $(\tilde{x}^\nu - x^\nu)/t^\nu$ converges to v, we know that ϵ^ν satisfies

$$\| \epsilon^\nu \| \leq c \frac{\| t^\nu \epsilon^\nu \|}{\| \tilde{x}^\nu - x^\nu \|},$$

for some positive constant $c < \infty$. The right side in the above inequality approaches zero since $G(\ ,\bar{p})$ is strictly differentiable at $\bar{x}$, so we conclude that the terms ϵ^ν also approach zero. If we define $\tilde{z}^\nu$ by

$$\tilde{z}^\nu := z^\nu + G(x^\nu, \bar{p}) - G(\bar{x}, \bar{p}) - \nabla_x G(\bar{x}, \bar{p})(x^\nu - \bar{x}),$$

we know by the strict differentiability of $G(\ ,\bar{p})$ that $\tilde{z}^\nu \to \bar{z}$. Moreover, $\tilde{z}^\nu + t^\nu(w^\nu + \epsilon^\nu)$ is contained in the set $G(\tilde{x}^\nu, \bar{p}) + N(\tilde{x}^\nu)$ and $\tilde{z}^\nu$ is contained in the set $G(x^\nu, \bar{p}) + N(x^\nu)$, which implies that $\tilde{x}^\nu$ is contained in $S(\tilde{z}^\nu + t^\nu(w^\nu + \epsilon^\nu), \bar{p})$ and that x^ν is in $S(\tilde{z}^\nu, \bar{p})$. Since $w^\nu + \epsilon^\nu \to w$, it follows that $v \in D_* S(\bar{z}, \bar{p} | \bar{x})(w)$ as claimed above. It also follows that the limit of the points

$$w^\nu + \epsilon^\nu - \frac{G(\tilde{x}^\nu, \bar{p}) - G(x^\nu, \bar{p})}{t^\nu}$$

is contained in the set $D_* N\big(\bar{x} | \bar{z} - G(\bar{x}, \bar{p})\big)(v)$. This gives the opposite inclusion to (6), which establishes the formula above and completes the proof.

REMARK. The proof of Lemma 4.1 works under more general assumptions than those made in this section; namely that the mapping $G(\ ,\bar{p})$ is strictly differentiable at $\bar{x}$. Note also that the proof of Lemma 4.1 actually shows that the strict derivative image set $D_* L_S(\bar{z}|\bar{x})(w)$ is always contained in the strict derivative of the mapping $z \mapsto S(z, \bar{p})$, which is generally a smaller set than the strict derivative of the mapping S.

We can now prove our main characterization of Lipschitzian localizations for solution multifunctions (4).

THEOREM 4.1. *For the solution mapping* S (4) *and its associated linearization multifunction* L_S (5), *and for* $(\bar{z}, \bar{p}, \bar{x}) \in \operatorname{gph} S$, *the following are equivalent*:

(a) S *has a Lipschitzian localization near* $(\bar{z}, \bar{p}, \bar{x})$.
(b) $D_* S(\bar{z}, \bar{p} | \bar{x})(0) = \{0\}$.
(c) $D_* L_S(\bar{z}|\bar{x})(0) = \{0\}$.
(d) L_S *has a Lipschitzian localization near* $(\bar{z}, \bar{x})$.
(e) $-\nabla_x G(\bar{x}, \bar{p})(v)$ *is an element of the set* $D_* N\big(\bar{x} | \bar{z} - G(\bar{x}, \bar{p})\big)(v)$ *only for* $v = 0$.

PROOF. The implication (a) $\Rightarrow$ (b) follows from Theorem 1.1, and both the implication (b) $\Rightarrow$ (c) as well as the equivalence between (c) and (e) follow from the relationships in Lemma 4.1. Theorem 1.3 gives the equivalence between (c) and (d) since L_S^{-1} is max hypomonotone near $(\bar{z}, \bar{x})$, and the equivalence of (d) and (a) follows from Levy and Poliquin (1997, Theorem 4.1).

Notice that Theorem 4.1 provides a characterization (between (a) and (b)) that is identical to that in our Theorem 1.2, yet the solution multifunction S above does not necessarily take points in one Euclidean space to sets in the same space. Of course, the linearization multifunction does have this property, so L_S serves as a bridge between the solution multifunction and the setting covered by Theorem 1.2.

We also note that condition (d) above was studied by Robinson (1980) in the case where the multifunction N represents the normal cone mapping associated with a convex set (in which case, the solution mapping S represents the solutions to a variational inequality). Robinson called this property "strong regularity," and he showed that the Lipschitzian localization property (a) was a consequence of strong regularity. These results were subsequently extended to solution multifunctions even more general than the ones we consider here; for example where the mapping N in (4) is assumed only to have locally closed graph (see Dontchev 1995a, b, Dontchev and Itag 1994a).

Dontchev and Rockafellar (1996) studied the subclass of solution multifunctions (4) where the multifunction N is the normal cone mapping associated with a polyhedral convex set. They characterized the existence of Lipschitzian localizations for these special solution multifunctions in terms of their Aubin continuity alone, without any conditions on the strict derivative. It is not known if Lipschitzian localizations of the more general solution multifunctions S (4) can be similarly characterized. Of course, our Theorem 4.1 provides this kind of "single condition" characterization, but in terms of the strict derivative, and without any reference to Aubin continuity.

Our Theorem 4.1 is thus important in two different ways: It gives new characterizations (in terms of the strict derivative—properties (b), (c), and (e)) of the local Lipschitzian behavior of solution mappings, and it expands the class of solution mappings whose local Lipschitzian behavior can be thus characterized. One example whose Lipschitzian stability cannot be characterized by previous work, is the solution mapping which gives the stationary points associated with a family of nonlinear programming problems. The nonlinear programming problems we consider are formulated in terms of smooth functions g_0 on $\mathbb{R}^n \times \mathbb{R}^d$ and g_i on $\mathbb{R}^n$ for $i = 1, \ldots, m$ as follows:

$$\text{(7)} \qquad \text{minimize} \quad g_0(x, p) - \langle z, x \rangle \quad \text{over all } x \in C,$$

where the set $C \subseteq \mathbb{R}^n$ is defined by

$$C := \{x : g_1(x) \le 0, \ldots, g_s(x) \le 0, \; g_{s+1}(x) = 0, \ldots, g_m(x) = 0\}$$

We concentrate our attention on points x that are *stationary points* for the minimization problem (7) in the sense of satisfying, in association with some multiplier vector y, the Karush-Kuhn-Tucker optimality conditions:

$$\text{(8)} \qquad \exists y = (y_1, \ldots, y_m) \in N_K\big(g_1(x), \ldots, g_m(x)\big) \quad \text{with}$$
$$\nabla_x g_0(x) + y_1 \nabla_x g_1(x) + \cdots + y_m \nabla_x g_m(x) = 0,$$

written here for convenience in terms of $N_K(u)$ denoting the set of normal vectors at u to the polyhedral cone

$$\text{(9)} \qquad K := \{u \in \mathbb{R}^m : u_1 \le 0, \ldots, u_s \le 0, u_{s+1} = 0, \ldots, u_m = 0\}$$

The Karush-Kuhn-Tucker conditions are of course necessary for a feasible solution $\bar{x}$ to be locally optimal under the Mangasarian-Fromovitz constraint qualification, which in turn takes the form

$$\text{(10)} \qquad \nexists y = (y_1, \ldots, y_m) \in N_K\big(g_1(\bar{x}), \ldots, g_m(\bar{x})\big) \quad \text{with}$$
$$y_1 \nabla_x g_1(\bar{x}) + \cdots + y_m \nabla_x g_m(\bar{x}) = 0, \quad \text{except } y = 0$$

Stationary points are sure to be optimal solutions when the minimization problem exhibits convexity with respect to x, but this is not an issue of concern to us here.

Under the Mangasarian-Fromovitz constraint qualification, the stationary points are exactly the solutions to the variational condition

$$z \in \nabla_x g_0(x, p) + N_C(x)$$

(see Levy and Rockafellar 1995a, for example), so that the solution mapping giving the stationary points for each parameter pair (z, p) is just

$$\text{(11)} \qquad S(z, p) = \{x : z \in \nabla_x g_0(x, p) + N_C(x)\}$$

Notice that the above variational condition is *not* necessarily a variational inequality (since the set C is not necessarily convex), nor is the set C polyhedral, so the characterization in Dontchev and Rockafellar (1996) does not apply in this case. However, the normal cone mapping N_C in (11) is max hypomonotone, so our Theorem 4.1 applies here to yield the following characterization.

COROLLARY 4.1. *If $\bar{x}$ is a stationary point of the nonlinear programming problem* (7) *for the parameters $\bar{z}$ and $\bar{p}$, and if the Mangasarian-Fromovitz constraint qualification holds at $\bar{x}$, then the stationary point mapping* (11) *has a Lipschitzian localization near $(\bar{z}, \bar{p})$ if and only if $v = 0$ is the only solution in $\mathbb{R}^n$ to*

$$-\nabla^2_{xx} g_0(\bar{x}, \bar{p})(v) \in D_* N_C\big(\bar{x} | \bar{z} - \nabla_x g_0(\bar{x}, \bar{p})\big)(v)$$

5. Karush Kuhn Tucker pairs. In the previous section, we showed that the stationary point mapping for a parameterized family of nonlinear programs is one example of a solution mapping whose sensitivity analysis needs the theory developed here. In this section, we study a solution mapping whose local Lipschitzian behavior has been characterized before, but we show how our results can be applied to this mapping to obtain a different characterization of its local Lipschitzian behavior. Thus, the previous section and this section together illustrate two advances that the theory in this paper provides; now a broader range of problems can be analyzed, and new insights can be gained into problems already within the range of the previous theory.

We consider a family of optimization problems very similar to the one in the previous section, but with an expanded parameterization. Now, all of the g_i for $i = 0, \ldots, m$ are smooth functions on $\mathbb{R}^n \times \mathbb{R}^d$, and the family of optimization problems becomes

$$\text{(12)} \qquad \text{minimize} \quad g_0(x, p) - \langle v, x \rangle \quad \text{over all } x \in C(u, p)$$

where the set $C(u, p) \subseteq \mathbb{R}^n$ is defined by

$$C(u, p) := \{x : g_1(x, p) + u_1 \le 0, \ldots, g_m(x, p) + u_m \le 0\}$$

(To simplify the calculations and without loss of generality, we consider only inequality constraints.) Here, the parameter pair $(v, u) \in \mathbb{R}^n \times \mathbb{R}^m$ plays a role like that of the canonical parameter z before, so we denote the pair by $z = (v, u)$.

Again, we focus on the Karush-Kuhn-Tucker optimality conditions, but this time as expressed in terms of

$$L(x, y, p) = g_0(x, p) + y_1 g_1(x, p) + \cdots + y_m g_m(x, p),$$

so that the KKT optimality conditions take the form

$$\text{(13)} \qquad \begin{aligned} v &= \nabla_x L(x, y, p) \\ u &\in -\nabla_y L(x, y, p) + N_{\mathbb{R}^m_+}(y), \end{aligned}$$

where $\mathbb{R}^m_+$ denotes the set of m-vectors with nonnegative components. The conditions (13) can be combined to yield the variational condition

$$z \in \big(\nabla_x L(x, y, p), -\nabla_y L(x, y, p)\big) + N_{\mathbb{R}^n \times \mathbb{R}^m_+}(x, y),$$

since the normal cone to $\mathbb{R}^n \times \mathbb{R}^m_+$ at (x, y) is just the set $\{0\} \times N_{\mathbb{R}^m_+}(y)$. Instead of studying the behavior of the stationary points as before, we now consider the pairs (x, y) satisfying the KKT optimality conditions, and we construct the following solution mapping:

$$\text{(14)} \qquad S(z, p) := \big\{(x, y) : z \in G(x, y, p) + N_{\mathbb{R}^n \times \mathbb{R}^m_+}(x, y)\big\},$$

where the mapping G is defined by

$$G(x,y,p) = \big(\nabla_x L(x,y,p), -\nabla_y L(x,y,p)\big)$$

According to Theorem 4.1, we can characterize the local Lipschitzian behavior of the solution mapping (14) in terms of the strict derivative of the mapping $N_{\mathbb{R}^n \times \mathbb{R}^m_+}$ (analogously to Corollary 4.1). With this in mind, we compute the strict derivative of this normal cone mapping. We rely on Lemma 2.1 and the fact that the graph of $N_{\mathbb{R}^n \times \mathbb{R}^m_+}$ is a kernel inverting Lipschitzian manifold for the kernel inverting matrix

$$A = \begin{bmatrix} 0 & I \\ I & -I \end{bmatrix},$$

and the Lipschitzian mapping $F: \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ defined by

$$F(u) = \left\{ v \,\middle|\, \begin{array}{ll} v_i = 0 & \text{for } i \in [n+1, m] \text{ with } u_i \le 0, \\ v_i = u_i & \text{else} \end{array} \right.$$

According to Lemma 2.1, the strict derivative of the normal cone mapping can be computed in terms of the strict derivative of F which has the form

$$D_* F(\bar{u})(u) = \left\{ v \,\middle|\, \begin{array}{ll} v_i = 0 & \text{for } i \in [n+1, m] \text{ with } \bar{u}_i < 0, \\ v_i \in [0, u_i] & \text{for } i \in [n+1, m] \text{ with } \bar{u}_i = 0 \le u_i, \\ v_i \in [u_i, 0] & \text{for } i \in [n+1, m] \text{ with } \bar{u}_i = 0 > u_i, \\ v_i = u_i & \text{else} \end{array} \right.$$

Lemma 2.1 then gives w in the set $D_* N_{\mathbb{R}^n \times \mathbb{R}^m_+}(\bar{x}, \bar{y} | \bar{z} - G(\bar{x}, \bar{y}, \bar{p}))(v)$ if and only if v is in the set $D_* F\big((\bar{x}, \bar{y}) + \bar{z} - G(\bar{x}, \bar{y}, \bar{p})\big)(v + w)$ which through the previous formula translates into

$$\begin{array}{ll} v_i = 0 & \text{for } i \in [n+1, m] \text{ with } \bar{z}_i - G(\bar{x}, \bar{y}, \bar{p})_i < -\bar{y}_i, \\ v_i\, w_i \ge 0 & \text{for } i \in [n+1, m] \text{ with } \bar{z}_i - G(\bar{x}, \bar{y}, \bar{p})_i = -\bar{y}_i, \\ w_i = 0 & \text{else} \end{array}$$

These computations and the definition of the mapping G lead directly to the following corollary to Theorem 4.1.

COROLLARY 5.1. *If* $(\bar{x}, \bar{y})$ *is a pair satisfying the KKT conditions* (13) *for the nonlinear programming problem* (12) *for the parameters* $\bar{z}$ *and* $\bar{p}$, *then the KKT-pair mapping* (14) *has a Lipschitzian localization near* $(\bar{z}, \bar{p})$ *if and only if* $(v, u) = 0$ *is the only solution in* $\mathbb{R}^n \times \mathbb{R}^m$ *to*

$$\begin{aligned} u_i &= 0 \quad \text{for } i \in I_1 := \{i : \bar{y}_i = 0 > \bar{u}_i + g_i(\bar{x}, \bar{p})\}, \\ u_i \langle \nabla_x g_i(\bar{x}, \bar{p}), v \rangle &\ge 0 \quad \text{for } i \in I_2 := \{i : \bar{y}_i = 0 = \bar{u}_i + g_i(\bar{x}, \bar{p})\}, \\ \langle \nabla_x g_i(\bar{x}, \bar{p}), v \rangle &= 0 \quad \text{for } i \in I_3 := \{i : \bar{y}_i > 0 = \bar{u}_i + g_i(\bar{x}, \bar{p})\} \\ v^T \big[\nabla^2_{xx} L(\bar{x}, \bar{y}, \bar{p})\big] &= -\sum_{I_2 \cup I_3} u_i \nabla_x g_i(\bar{x}, \bar{p}) \end{aligned}$$

Many other characterizations of this sort have been made previously for the KKT-pair mapping (see for example, Dontchev and Rockafellar 1996, Kojima 1988, Kummer 1991, Klatte and Tammer 1990), but the characterization we present in Corollary 5.1 is different from any of these. Some of the previous results have been framed in terms of the strong second-order condition

(15) $\nabla^2_{xx} L(\bar{x}, \bar{y}, \bar{p})$ is positive definite on $\{v : v \perp \nabla_x g_i(\bar{x}, \bar{p}) \text{ for } i \in I_3\}$,

and the linear independence of the constraint gradients for indices in the set $I_2 \cup I_3$. Note that the latter of these conditions follows immediately from the characterization in Corollary 5.1: If there is a combination of u_i such that

$$\sum_{I_2 \cup I_3} u_i \nabla_x g_i(\bar{x}, \bar{p}) = 0,$$

then it follows that the pair $(0, u) \in \mathbb{R}^n \times \mathbb{R}^m$ (where u is augmented with zeroes in the components corresponding to indices in I_1) satisfies all four conditions in Corollary 5.1. Note also that the strong second-order condition (15) together with the linear independence of the constraint gradients for indices in $I_2 \cup I_3$, is a sufficient condition for the zero vector to be the only solution to the four conditions in Corollary 5.1 (multiply both sides of the last equation by the vector v and apply the first three conditions). In fact, it has been shown previously that the strong second-order condition (15) paired with the linear independence of the constraint gradients for indices in $I_2 \cup I_3$ is indeed equivalent to the existence of a Lipschitzian localization of the KKT-pair mapping for which the primal elements are actually locally optimal (see Dontchev and Rockafellar 1996, Theorem 6 or Robinson 1980 and Bonnans and Sulem 1995).

References

Bonnans, J. F., A. Sulem, 1995. Pseudopower expansion of solutions of generalized equations and constrained optimization problems. *Math. Programming* **70** 123–148.

Dontchev, A. L. 1995a. Characterizations of Lipschitz stability in optimization. *Recent Developments in Well-posed Variational Problems*. Kluwer, 95–116.

——. 1995b. Implicit function theorems for generalized equations. *Math. Programming* **70** 91–106.

——, W. W. Hager, 1994a. Implicit functions, Lipschitz maps, and stability in optimization. *Math. Oper. Res.* **19** 753–768.

——, W. W. Hager, 1994b. An inverse mapping theorem for set-valued maps. *Proc. Amer. Math. Soc.* **121** 481–489.

——, R. T. Rockafellar. 1996. Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM J. Optim.* **6** 1087–1105.

Klatte, D., K. Tammer. 1990. Strong stability of stationary solutions and Karush-Kuhn-Tucker points in nonlinear optimization. *Ann. Oper. Res.* **27** 285–308.

Kojima, M. 1988. Strongly stable stationary solutions in nonlinear programming. S. M. Robinson, ed., *Analysis and Computation of Fixed Points*. Academic Press, 93–138.

Kummer, B. 1991. Lipschitzian inverse functions, directional derivatives, and applications in $\mathscr{C}^{1,1}$ optimization. *J. Optim. Theory Appl.* **70**.

Levy, A. B., R. A. Poliquin. 1997. Characterizing the single-valuedness of multifunctions. *Set-Valued Anal.* **5** 351–364.

——, R. T. Rockafellar. 1995a. Sensitivity of solutions in nonlinear programming problems with nonunique multipliers. D.-Z. Du, L. Qi, and R. S. Womersley, Eds. *Recent Advances in Nonsmooth Optimization*. World Scientific, 215–223.

——, ——. 1995b. Variational conditions and the proto-differentiation of partial subgradient mappings. *Nonlinear Anal. Theory, Methods, Appl.* **26** 1951–1964.

——, ——. 1996. Proto-derivatives and the geometry of solution mappings in nonlinear programming. *Nonlinear Optimization and Applications*, Plenum, 343–365.

Poliquin, R. A. 1991. Integration of subdifferentials of nonconvex functions. *Nonlinear Anal., Theory, Methods and Appl.* **17** 385–398.

——, R. T. Rockafellar. 1995. Prox-regular functions in variational analysis. *Trans. Amer. Math. Soc.* **348** 1805–1838.

Robinson, S. M. 1980. Strongly regular generalized equations. *Math. Oper. Res.* **5** 43–62.

Rockafellar, R. T. 1985. Maximal monotone relations and the second derivatives of nonsmooth functions. *Ann. l'Institût Henri Poincaré—Anal. non linèaire* **2** 167–184.

——, R. J.-B. Wets. 1998. *Variational Analysis*. Springer Verlag,

A. B. Levy: Department of Mathematics, Bowdoin College, Brunswick, Maine 04011