LIPSCHITZIAN STABILITY OF NEWTON'S METHOD FOR VARIATIONAL INCLUSIONS*

Asen L. Dontchev
AMS, Mathematical Reviews
Ann Arbor, MI 48107-8604, USA
ald@ams.org

Abstract We present an overview to Lipschitz-type properties of mappings associated with solutions of optimization problems including variational inequalities and mathematical programs. We show that these properties are inherited in various ways by the mapping acting from parameters of the problem and the starting point to the set of sequences generated by Newton's method. Some new insights into convergence of Newton's/SQP method are also presented.

Keywords: Newton's method, stability, variational inequalities, optimization.

1. INTRODUCTION

In this paper we study continuity properties of solutions to variational problems and associated properties of sequences generated by Newton's method. Our basic model is the following inclusion (generalized equation) with a parameter v,

$$v \in f(x) + F(x), \tag{1}$$

where $v \in \mathbf{R}^m$, $x \in \mathbf{R}^n$ f is a C^2 function from $\mathbf{R}^n$ to $\mathbf{R}^m$ and F is a set-valued map from $\mathbf{R}^n$ to the subsets of $\mathbf{R}^m$ with closed graph. For $F(x) = \{0\}$ we obtain a system of equations, for $F(x) = \mathbf{R}^m_+$, the positive orthant in $\mathbf{R}^m$, we have a system of inequalities. If $F(x) = N_C(x)$, the normal cone mapping to a closed set C in $\mathbf{R}^n$, for $m = n$, then we have a variational inequality; in particular, for $C = \mathbf{R}^n_+$ we obtain a complementarity problem.

*This research was supported by the National Science Foundation.

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: 10.1007/978-0-387-35514-6_15

M.J.D. Powell and S. Scholtes (Eds.), *System Modelling and Optimization: Methods, Theory and Applications.*

The parameter v provides a *shift* perturbation to the inclusion (1). Most of the results presented in this paper can be extended to problems with a *basic* perturbation parameter, u, added to the function f and to the mapping F, thus regarding (1) as embedded in a larger family of inclusions,

$$v \in f(x, u) + F(x, u).$$

The pair (v, u) now represents *canonical* perturbations. In general, shift stability of (1) with respect to v is different from full (canonical) stability with respect to (u, v); in some cases these concepts are equivalent under additional conditions for the dependence of f and F on the "generic" parameter u. In this paper we stay within the format of shift stability of the model (1). We study local Lipschitz-type properties of the map "parameter $v \mapsto$ set of solutions of (1)", denoted $\mathbf{S}$; that is,

$$v \mapsto \mathbf{S}(v) = \{x \in \mathbf{R}^n \mid v \in f(x) + F(x)\}, \tag{2}$$

around a fixed reference point (v^*, x^*) in the graph of $\mathbf{S}$. Of course, $\mathbf{S}$ is the inverse map $(f + F)^{-1}$.

For illustration of our results we consider throughout the following optimization problem:

$$\min_x \, [g(x) - \langle v, x \rangle] \quad \text{subject to} \;\; x \in C, \tag{3}$$

where g is a $\mathcal{C}^3$ function from $\mathbf{R}^n$ to $\mathbf{R}$, C is a convex polyhedral set in $\mathbf{R}^n$ and v is a parameter which provides "tilting" perturbations to the cost, in the terminology of [17]. The first-order necessary optimality condition for (3) has the form

$$v \in \nabla g(x) + N_C(x), \tag{4}$$

where ∇g is the derivative of g and $N_C(x)$ denotes the normal cone to the set C at the point x. The variational inequality (4) fits into the general format of (1).

By Newton's method applied to the basic model (1) we mean the procedure which generates a sequence $\{x_1, x_2, \cdots\}$ of points x_n, with a given x_0, according to the rule

$$v \in f(x_n) + \nabla f(x_n)(x_{n+1} - x_n) + F(x_{n+1}) \quad \text{for } n = 0, 1, \cdots. \tag{5}$$

This procedure is the standard Newton for equations when $F = \{0\}$; it is usually called a generalized Newton method or a Newton-type method for solving systems of inequalities when $F = \mathbf{R}^m_+$ or variational inequalities when F is a normal cone mapping, respectively. One may expect that the "true" Newton method for the general model (1) should involve some kind of linearization of the map F, and a number of authors have followed this line of research for various nonsmooth, semismooth, etc. equations, see, e.g., the recent survey in [26]. Here we keep the map F unchanged with the mental picture that F has a "simple" yet set-valued form, e.g. it is a polyhedral map or the normal map to a polyhedral cone, both having already a linear structure. Such a model leads us quite far, also for infinite-dimensional (i.e., optimal control) problems where nonsmooth-analytic tools may become more technical than instrumental. Most of the results in this paper hold in abstract spaces, by changing only the terminology.

Consider the optimization problem (3). Representing the convex polyhedral set C as $C = \{x \in \mathbf{R}^n \mid Bx \leq b\}$ for some matrix B from $\mathbf{R}^n$ to $\mathbf{R}^m$ and a vector $b \in \mathbf{R}^m$, the normal cone to the set C at x has the form

$$N_C(x) = \{v \in \mathbf{R}^n \mid v = B^T y \text{ for some } y \in \mathbf{R}^m_+ \text{ and } x \perp v\}.$$

Then the Newton method (5) applied to the variational inequality (4) becomes

$$\begin{aligned} &-v + \nabla g(x_n) + \nabla^2 g(x_n)(x_{n+1} - x_n) + B^T y_{n+1} = 0, \\ &Bx_{n+1} \leq b, \quad y_{n+1} \geq 0, \quad (Bx_{n+1} - b)^T y_{n+1} = 0. \end{aligned} \tag{6}$$

This is the best-known form of the sequential quadratic programming (SQP) method for the problem (3); indeed, an iteration in (6) is obtained to solving a quadratic programming problem.

For given x and v, let $\xi = \{x_1, x_2, \cdots, x_n, \cdots\}$ be a Newton sequence, that is, a sequence starting from x and satisfying (5). Denote by $\mathbf{N}(x, v)$ the set of all Newton sequences, starting from the point x for v. Let $\xi^* = \{x^*, x^*, \cdots, x^*, \cdots\}$, that is, ξ^* is the constant sequence with all elements x^*. Note that $\xi^* \in \mathbf{N}(x^*, v^*)$. We equip the set of Newton sequences with a distance induced by the l^∞ norm:

$$\|\xi\|_\infty = \sup_{n\geq 1} \|x_n\|.$$

In this paper we study continuity properties of the map $(x, v) \mapsto \mathbf{N}(x, v)$ in relation to corresponding properties of the map $\mathbf{S}$ defined in (2). We consider three continuity properties which play central roles in quantitative stability analysis of variational problems: the local upper-Lipschitz property, the Aubin continuity, and the Lipschitzian localization property. We also provide some new insights into convergence properties of Newton's method.

In this paper we apply the basic principle of smooth nonlinear analysis, stated by A. Ioffe [10] in the following way: a property of the derivative of a nonlinear map at a given point should hold, in a sufficiently small neighborhood, for the map itself. In the context of the inclusion (1) and the corresponding solution map $\mathbf{S} = (f + F)^{-1}$, this principle means that adding terms of order $o(x)$ to (1) doesn't change certain properties of the map **S**. Since replacing f by its linearization at x^* corresponds to adding $o(x - x^*)$, we call this specific form of the basic principle *invariance under linearization.*

The basic principle of smooth nonlinear analysis is present already in the classical implicit function theorem and is even more explicit in the Lyusternik and Graves theorems. In his seminal paper [21], Robinson extended this principle to variational inequalities and optimization problems. The main idea in the present paper is to apply the basic principle to Lipschitz-type properties of Newton's map **N** defined above.

In Section 2 we study the local upper-Lipschitz continuity, a relatively recently introduced concept which turns out to be quite natural in nonlinear optimization. We present characterizations of the upper-Lipschitz continuity for the map **S** and for the Argmin map of the problem (3). In particular, we show that the local upper-Lipschitz continuity is invariant under linearization. We prove that the Argmin map of the problem (3) possesses this property if and only if the standard second-order sufficient optimality condition holds.

In Section 3 we first show that the local upper-Lipschitz continuity of the map **S** implies that every Newton sequence within a sufficiently small ball around the solution is quadratically convergent. Then we study the local upper-Lipschitz continuity of the Newton map **N**. We prove that if the map **S** is locally upper-Lipschitz, then the map **N** is locally upper-Lipschitz as well (in the l^∞ norm). We also show that the converse implication holds if, in addition, **N** is locally nonempty-valued. For the optimization problem (3) this result evolves into the equivalence between the second-order sufficient optimality condition and the property that the Newton (SQP) map is locally upper-Lipschitz and nonempty-valued.

At the end of Section 3 we briefly show that analogous results can be obtained for the proximal point method.

In Section 4 we discuss the Aubin continuity by following the pattern of the previous sections. We give a version of the Lyusternik-Graves theorem (Lemma 4.1) and use it to show that the Aubin continuity of the map **S** implies existence of locally quadratically convergent Newton sequences around the reference point. Further, the Newton map **N** is Aubin continuous if the map **S** is Aubin continuous. As a partial converse, the Aubin continuity of the submap of **N** associated with all convergent Newton sequences implies the Aubin continuity of the map **S**.

Section 5 is devoted to the Lipschitzian localization, a continuity property which appears in implicit function theorems and is best studied in optimization. We show that the Lipschitzian localization of the map $\mathbf{S}$ implies the existence of a unique Newton sequence around the reference point. Also, the Newton map $\mathbf{N}$ has a Lipschitzian localization if and only if the map $\mathbf{S}$ possesses this property. For the problem (3) the Lipschitzian localization of the Argmin map is equivalent to the strong second-order sufficient optimality condition.

There is a vast literature on the continuity concepts discussed in the present paper and their applications to various variational problems. For recent discussions and references see the bibliographical comments in [10, 12, 25].

As a further application of the approach presented in this paper, we target error analysis of discrete approximations to infinite-dimensional variational problems for which the parameter v represents the discretization remainder while $v = v^*$ is identified with the original (continuous) problem. We shall not discuss here discrete approximations; for first results in this direction, see the forthcoming paper [6].

Throughout we denote by $\|\cdot\|$ any norm in $\mathbf{R}^n$, by $B_r(x)$ the closed ball with center x and radius r, and by $\mathbf{B}$ the ball $B_1(0)$. In writing "f maps X into Y" we mean that the domain of f is a (possibly proper) subset of X; thus, a set-valued map F from X to the subsets of Y may have empty values for some points of X. Given a map F from X to the subsets of Y, we define $\operatorname{graph} F = \{(x, y) \in X \times Y \mid y \in F(x)\}$ and $F^{-1}(y) = \{x \in X \mid y \in F(x)\}$. We denote by $\operatorname{dist}(x, A)$ the distance from a point x to a set A and by ^{T} a transposition of a matrix or a vector. Throughout L is a Lipschitz constant of $\nabla f(\cdot)$ in a (sufficiently large) ball centered at x^*.

2. LOCAL UPPER-LIPSCHITZ CONTINUITY

The local upper-Lipschitz continuity is a localized version, for the graph of a map, of the upper-Lipschitz continuity introduced by Robinson [20].

Definition 2.1 *A map $\Gamma : \mathbf{R}^n \to \mathbf{R}^m$ is locally upper-Lipschitz continuous at $(y^*, x^*) \in \operatorname{graph}\Gamma$ with constants a and b for neighborhoods and c for growth if*

$$\Gamma(y) \cap B_a(x^*) \subset \{x^*\} + c\|y - y^*\|\mathbf{B} \quad \textit{for all } \; y \in B_b(y^*).$$

The following properties can be deduced directly from the definition: A locally upper-Lipschitz map Γ at (x^*, y^*) has necessarily x^* as a locally unique (isolated) point of $\Gamma(y^*)$. If a map is locally upper-Lipschitz with constants a, b and c, then for every $0 \le a' \le a, 0 \le b' \le b$, and $c' \ge c$, it is locally upper-Lipschitz with constants a', b' and c'. If a map Γ is locally upper-Lipschitz at (y^*, x^*) and Σ is a local selection of Γ around (x^*, y^*), that is, for some neighborhoods U of x^* and V of y^* we have $\Sigma(y) \cap U \subset \Gamma(y) \cap U$ for $y \in V$, then Σ is locally upper-Lipschitz at (y^*, x^*).

Robinson [22] proved that in finite dimensions every map whose graph is a polyhedral (possibly nonconvex) set is upper-Lipschitz at every point of its domain. As a consequence, every map $\Gamma : \mathbf{R}^n \to \mathbf{R}^m$, for which x^* is an isolated point of $\Gamma(y^*)$ and $\operatorname{graph}\Gamma$ is a polyhedral set, is locally upper-Lipschitz at (y^*, x^*). Such maps are for instance the solution maps of the linear variational inequalities over convex polyhedral sets when the reference solution is isolated. We note that Robinson used in [20] a different definition of the local upper-Lipschitz continuity where the localization is in the domain of the map.

Rockafellar [23], see also [11, 15], gave a characterization of the upper-Lipschitz property by employing the so-called proto-derivatives (contingent derivatives). We showed in [4] that the local upper-Lipschitz continuity is invariant under linearization in a very abstract setting (see Lemma 2.1 and Corollary 2.1 below). The upper-Lipschitz property was studied in [8] for a general nonlinear programming problem with canonical parameters in both the functional and the constraints. In [8], Theorem 2.6, it was proved that the Karush-Kuhn-Tucker map has this property with the reference point being a local optimal solution exactly when the combination of the strict Mangasarian-Fromowitz condition and the standard second-order sufficient optimality condition holds. Further extensions and generalizations of this result for nonsmooth mathematical programs, as well as extended surveys on the subject, are given in the recent papers [9, 12, 16].

In a related vain, Zolezzi introduced and studied in [27] a condition number for the solutions to abstract optimization problems of the form $\min f(x, y)$, $x \in X$, where y is a parameter. This conditional number is precisely the growth constants c in the definition of the upper-Lipschitz property, where $\Gamma(y)$ is the set of optimal solutions for y.

Note that the local upper-Lipschitz property does not guarantee nonemptiness of the values of Γ near the reference point y^* (as an example, take $\Gamma(y^*) = x^*$ and $\Gamma(y) = \emptyset$ for $y \ne y^*$). Such a requirement leads to a different concept, namely:

Definition 2.2 *A map* $\Gamma : \mathbf{R}^n \to \mathbf{R}^m$ *is locally nonempty-valued at* $(y^*, x^*) \in \operatorname{graph}\Gamma$ *if there exist positive numbers* a *and* b *such that*

$$\Gamma(y) \cap B_a(x^*) \ne \emptyset \quad \textit{for all } \; y \in B_b(y^*).$$

A map Γ is locally nonempty-valued at (y^*, x^*) if and only if the map Γ^{-1} is open at (x^*, y^*). If $\Gamma(y)$ is the set of solutions of a system of inequalities with data y (matrices, vectors), then the radius of the neighborhood b is a kind of bound on perturbations of the data from y^* such that the system still has a solution x within distance a from the reference solution x^* for y^*. With $a = \infty$ and after normalization, the supremum of such b is exactly the relative condition number introduced by Renegar [19].

The following lemma shows the invariance under linearization of the upper-Lipschitz continuity in a transparent form which is particularly suitable for applications:

Lemma 2.1 *Let G be a map from $\mathbf{R}^n$ to the subsets of $\mathbf{R}^m$ and let G^{-1} be locally upper-Lipschitz at (v^*, x^*) with constants a and b for neighborhoods and c for growth. Let the nonnegative constants α, β and λ satisfy*

$$\alpha \leq a, \quad \lambda c < 1, \quad \beta + \lambda\alpha \leq b. \tag{7}$$

Let $h : \mathbf{R}^n \to \mathbf{R}^m$ be a function which satisfies

$$\|h(x) - h(x^*)\| \leq \lambda\|x - x^*\| \quad \text{for all } x \in B_\alpha(x^*). \tag{8}$$

Then the map $(h + G)^{-1}$ is locally upper-Lipschitz at $(v^* + h(x^*), x^*)$ with constants α and β for neighborhoods and $c/(1 - \lambda c)$ for growth.

Proof. Choose α, β and λ as in (7) and let h satisfy (8). Put $y^* = v^* + h(x^*)$ and let $y \in B_\beta(y^*)$ and $x \in (h + G)^{-1}(y) \cap B_\alpha(x^*)$. Then $x \in G^{-1}(y - h(x)) \cap B_a(x^*)$ and

$$\|y - h(x) - v^*\| \leq \|y - y^*\| + \|h(x) - h(x^*)\| \leq \beta + \lambda\alpha \leq b.$$

Hence, from the upper-Lipschitz continuity of G^{-1} at (v^*, x^*) we obtain

$$\begin{aligned} \|x - x^*\| \leq c\|y - h(x) - v^*\| &\leq c\|y - y^*\| + c\|h(x) - h(x^*)\| \\ &\leq c\|y - y^*\| + c\lambda\|x - x^*\|. \end{aligned}$$

Thus,

$$\|x - x^*\| \leq \frac{c}{1 - \lambda c}\|y - y^*\|$$

and the proof is complete. □

It is clear from the proof that this lemma still holds if we replace $\mathbf{R}^n$ with any metric space and $\mathbf{R}^m$ with a linear space with a shift-invariant metric.

Recall that $\mathbf{S}(v)$ denotes the set of solution to (1) for v. With the solution map $\mathbf{S}$ we associate the map

$$v \mapsto \mathbf{L}(v) = (f(x^*) + \nabla f(x^*)(\cdot - x^*) + F(\cdot))^{-1}(v)$$

obtained by the linearization of f at the point x^*. We have $x^* \in \mathbf{L}(v^*)$ iff $x^* \in \mathbf{S}(v^*)$. From Lemma 2.1 we obtain the following result for the solutions of (1) first stated in [4], Theorem 3.2.

Corollary 2.1 *The following are equivalent:*

(i) The map $\mathbf{L}$ *is locally upper-Lipschitz at* (v^*, x^*).

(ii) The map $\mathbf{S}$ *is locally upper-Lipschitz at* (v^*, x^*).

We note that the local nonempty-valuedness of a map (Definition 2.2) is not invariant under linearization.

In our analysis of Newton's method we use the following map:

$$(v, x) \mapsto \mathbf{P}(v, x) := (f(x) + \nabla f(x)(\cdot - x) + F(\cdot))^{-1}(v). \tag{9}$$

If x^* is a solution to (1) for v^*, then $x^* \in \mathbf{P}(v^*, x^*)$. Also, $\mathbf{L}(v) = \mathbf{P}(v, x^*)$. The following corollary complements Corollary 2.1.

Corollary 2.2 *The following are equivalent:*

(i) The map $\mathbf{L}$ *is locally upper-Lipschitz at* (v^*, x^*).

(ii) The map $\mathbf{P}$ *is locally upper-Lipschitz at* $((v^*, x^*), x^*)$.

Proof. The implication (ii) $\Rightarrow$ (i) is immediate by noting that $\mathbf{L}(v) = \mathbf{P}(v, x^*)$. Let us prove (i) $\Rightarrow$ (ii). Let $\mathbf{L}$ be locally upper-Lipschitz at (v^*, x^*) with constants a, b and c and let the positive numbers α and β satisfy the inequalities (7) with $\lambda = L\beta$. (Recall that, here and later, L is a Lipschitz constant of $\nabla f(\cdot)$ in a (sufficiently large) ball centered at x^*; in this case this is the ball with radius a.)

Let $(v, x) \in B_\beta((v^*, x^*))$ and let $z \in \mathbf{P}(v, x) \cap B_a(x^*)$. We apply Lemma 2.1 with

$$G(z) = f(x^*) + \nabla f(x^*)(z - x^*) + F(z)$$

and

$$h(z) = f(x) + \nabla f(x)(z - x) - f(x^*) - \nabla f(x^*)(z - x^*).$$

For any $z' \in B_\alpha(x^*)$ we have

$$\begin{aligned} \|h(z') - h(x^*)\| &= \|(\nabla f(x) - \nabla f(x^*))(z' - x^*)\| \\ &\leq L\beta\|z' - x^*\| = \lambda\|z' - x^*\|, \end{aligned}$$

hence (8) holds. By assumption, G^{-1} is locally upper-Lipschitz at (v^*, x^*), hence from Lemma 2.1 we obtain that the map $(h+G)^{-1}$ is locally upper-Lipschitz at $(v^* + h(x^*), x^*)$. Note that $z \in (h+G)^{-1}(v) \cap B_\alpha(x^*)$. Then, for $\gamma = c/(1-\lambda c)$, we have

$$\begin{aligned} \|z - x^*\| &\leq \gamma\|v - v^* - f(x) - \nabla f(x)(x^* - x) + f(x^*)\| \\ &\leq \gamma\|v - v^*\| + \frac{1}{2}\gamma L\beta\|x - x^*\|. \end{aligned}$$

Thus the map **P** is locally upper-Lipschitz with constants α and β for the radia of the neighborhoods and growth constants γ for v and $\gamma L\beta/2$ for x, respectively. □

Remark 2.1 *Note that the growth constant of* **P** *with respect to x can be made arbitrarily small, by choosing β small.*

Consider the optimization problem (3) and the corresponding first-order necessary optimality condition (4). Let x^* be a solution of (4) for $v = v^*$. Denote by K the critical cone at (v^*, x^*), that is,

$$K = \{x \in T_C(x^*) \mid v^* - \nabla g(x^*) \perp x\},$$

where $T_C(x^*)$ is the tangent cone to C at the point x^*. Then the (standard) second-order sufficient optimality condition at (v^*, x^*) has the form

$$\langle u, \nabla^2 g(x^*)u\rangle > 0 \quad \text{for all nonzero } u \in K. \tag{10}$$

Let Argmin(v) be the set of local solutions for v to the optimization problem (3).

Theorem 2.1 *The following are equivalent:*

(i) The second-order sufficient optimality condition (10) *holds at* (v^*, x^*).

(ii) The Argmin map is locally nonempty-valued and upper-Lipschitz continuous at (v^*, x^*).

Proof. Let $\mathcal{X}$ be the map of the critical points,

$$v \mapsto \mathcal{X}(v) = \{x \in \mathbf{R}^n \mid v \in \nabla g(x) + N_C(x)\}. \tag{11}$$

The solution map of the linearization of (4) at x^* is defined as

$$v \mapsto \mathcal{L}(v) = \{x \in \mathbf{R}^n \mid v \in \nabla g(x^*) + \nabla^2 g(x^*)(x - x^*) + N_C(x)\}. \tag{12}$$

Let (i) hold. For short, denote $a = \nabla g(x^*)$ and $A = \nabla^2 g(x^*)$. First, let us show that x^* is an isolated solution of the linear variational inequality

$$v^* \in a + A(x - x^*) + N_C(x). \tag{13}$$

On the contrary, suppose that there exists a sequence $x_n \to x^*$ such that

$$v^* \in a + A(x_n - x^*) + N_C(x_n) \quad \text{for all } n.$$

Then

$$v^* \in a + \nabla^2 g(x_n)(x_n - x^*) + o(\|x_n - x\|) + N_C(x^* + (x_n - x^*)).$$

By Reduction Lemma in [7],

$$0 \in \nabla^2 g(x_n)(x_n - x^*) + o(\|x_n - x^*\|) + N_K(x_n - x^*);$$

that is,

$$\langle \nabla^2 g(x_n)(x_n - x^*), x_n - x^* \rangle + o(\|x_n - x^*\|^2) = 0. \tag{14}$$

Since

$$b_n := \frac{x_n - x^*}{\|x_n - x^*\|} \in K \quad \text{for all } n \quad \text{and} \quad \|b_n\| = 1,$$

we obtain that a subsequence of b_n is convergent to, say, $u \in K$. From (14) we get

$$\langle \nabla^2 g(x_n) b_n, b_n \rangle + O(\|x_n - x^*\|) = 0.$$

Passing to the limit with (the subsequence of) $n \to \infty$ in the latter equality we come to a contradiction with (i). Thus x^* is an isolated solution of (13).

As noted at the beginning of this section, since $\mathcal{L}$ is a polyhedral map, from a result of Robinson [22] we obtain that the requirement x^* be an isolated point of $\mathcal{L}(v^*)$ is equivalent to the local upper-Lipschitz continuity of $\mathcal{L}$ at (v^*, x^*). Then the local upper-Lipschitz continuity of the critical point map $\mathcal{X}$ at (v^*, x^*) follows from the invariance under linearization (Corollary 2.1).

Consider the problem (3) with the additional constraint $\|x - x^*\| \le \delta$, for $\delta > 0$. By the Berge theorem, the solution map Argmin_δ of the new problem is upper semicontinuous at $v = v^*$. Clearly, Argmin_δ is nonempty-valued and $\text{Argmin}_\delta(v^*) = \{x^*\}$ for δ sufficiently small. One then sees that for v close to v^* eventually the constraint $\|x - x^*\| \le \delta$ is not active so that $\text{Argmin}_\delta(v) = \text{Argmin}(v) \cap B_\delta(x^*)$. Thus, Argmin is locally nonempty-valued at (v^*, x^*). Since $\text{Argmin}(v) \subset \mathcal{X}(v)$, we obtain that (ii) holds.

Conversely, let Argmin be locally nonempty-valued and upper-Lipschitz continuous at (v^*, x^*) with a growth constant κ and neighborhoods U and V. Without loss of generality, assume $v^* = 0$. Take any $x \in \mathbf{R}^n$ close x^* such that $v = (x - x^*)/2\kappa \in V$. Let $x_v \in \text{Argmin}(v) \cap U$ be an associated local minimizer, then $\|x_v - x^*\| \leq \kappa\|v\|$. From the optimality we have

$$g(x) - \langle v, x\rangle \geq g(x_v) - \langle v, x_v\rangle,$$

that is

$$g(x) - \frac{1}{2\kappa}\langle x - x^*, x - x^*\rangle \geq g(x_v) - \frac{1}{2\kappa}\langle x - x^*, x_v - x^*\rangle.$$

Since $g(x_v) \geq g(x^*)$ (remember that $v^* = 0$), we continue the chain of inequalities obtaining

$$\begin{aligned} g(x) - \frac{1}{2\kappa}\langle x - x^*, x - x^*\rangle &\geq g(x^*) - \frac{1}{2\kappa}\|x - x^*\|\|x_v - x^*\| \\ &\geq g(x^*) - \frac{1}{2\kappa}\|x - x^*\|\kappa\|v\| \\ &= g(x^*) - \frac{1}{4\kappa}\|x - x^*\|^2. \end{aligned}$$

Thus

$$g(x) \geq g(x^*) + \frac{1}{4\kappa}\|x - x^*\|^2$$

for every x close to x^*. We obtain the so-called growth condition of order two which, as well known, see e.g. [25], p. 606, is equivalent to the second-order sufficient condition (10). □

The above result clarifies the meaning of the second-order sufficient condition: it gives not only optimality at the reference point but also a certain continuity property of the minimizers with respect to certain perturbations and this property is exactly captured by the local upper-Lipschitz continuity.

The approach presented in this section can be applied to other mappings in variational analysis. As an example consider the map

$$(u, v) \mapsto \mathbf{M}_\mu(u, v) = (f(\cdot) + \mu(\cdot - u) + F(\cdot))^{-1}(v), \tag{15}$$

where μ is a positive scalar. In the context of our illustrative problem (3), the map $\mathbf{M}_\mu$ is the stationary point map of the prox-regularization (Moreau-Yosida regularization) of (3):

$$\min_{x \in C}\left[g(x) - \langle v, x\rangle + \frac{1}{2}\mu\|x - u\|^2\right].$$

A straightforward application of Lemma 2.1 gives us

Corollary 2.3 *Let the map* $\mathbf{S}$ *be locally upper-Lipschitz at* (v^*, x^*) *with a growth constant* c. *Then for every positive* $\mu < 1/c$ *the map* $\mathbf{M}_\mu$ *is locally upper-Lipschitz at* $((v^*, x^*), x^*)$ *with growth constants* $c/(1 - c\mu)$ *for* v *and* $c\mu/(1 - c\mu)$ *for* x.

At the end of the following section we apply this result to show convergence of the proximal-point method.

3. LOCAL UPPER-LIPSCHITZ CONTINUITY OF THE NEWTON MAP

Consider the following perturbed version of the Newton method (5) in which $\nabla f(x_n)$ is replaced by a sequence of matrices H_n and the parameter v may vary from step to step:

$$v_n \in f(x_n) + H_n(x_{n+1} - x_n) + F(x_{n+1}) \quad \text{for } n = 0, 1, \cdots. \tag{16}$$

Our first result relates the local upper-Lipschitz property of the map $\mathbf{L}$ (or, equivalently, $\mathbf{S}$) to convergence of Newton sequences.

Theorem 3.1 *Let* $\mathbf{L}$ *be locally upper-Lipschitz at* (v^*, x^*) *with constants* a *and* b *for neighborhoods and* c *for growth. Let the positive constants* σ *and* κ *satisfy*

$$\sigma \leq a, \quad c\kappa < 1, \quad \kappa(1 + 2\sigma) + \frac{1}{2}L\sigma^2 \leq b, \tag{17}$$

let $\{v_n\}$ *be a sequence of vectors with* $v_n \in B_\kappa(v^*), n = 1, 2, \cdots$, *and let* H_n *be a sequence of matrices with* $H_n \in B_\kappa(\nabla f(x^*))$, $n = 0, 1, \cdots$. *Then for every initial point* $x_0 \in B_\sigma(x^*)$, *every sequence* $\{x_n\}$ *obtained from (16) and whose elements are all in* $B_\sigma(x^*)$ *for* $n = 1, 2, \cdots$ *satisfies*

$$\|x_{n+1} - x^*\| \leq \frac{c}{1 - c\kappa}\Big(\|x_n - x^*\|^2 + \|(H_n - \nabla f(x^*))(x_n - x^*)\| + \|v_n - v^*\|\Big). \tag{18}$$

Proof. Choose κ and σ such that the inequalities (17) hold and let $\{x_n\}$ be a sequence satisfying the conditions of the theorem. Fix $n \geq 1$. We apply Lemma 2.1 with $G = \mathbf{L}^{-1}$, $h(x) = f(x_n) + H_n(x - x_n) - f(x^*) - \nabla f(x^*)(x - x^*)$, $\alpha = \sigma$, $\beta = \kappa(1 + \sigma) + \frac{1}{2}L\sigma^2, \lambda = \kappa$. The so defined

α, β and λ satisfy (7) and h satisfies (8). From Lemma 2.1, the map $(h+G)^{-1}$ is locally upper-Lipschitz at $(v^* + h(x^*), x^*)$ with constants α and β for the neighborhoods and a growth constant $c(1-c\lambda)$. We have

$$x_{n+1} \in (h+G)^{-1}(v_n) \cap B_\sigma(x^*)$$

and

$$\begin{aligned} &\|v_n - v^* - h(x^*)\| \le \|v_n - v^*\| + \|f(x_n) + H_n(x^* - x_n) - f(x^*)\| \\ &\le \|v_n - v^*\| + \|(H_n - \nabla f(x^*))(x_n - x^*)\| + \frac{1}{2}L\|x_n - x^*\|^2 \\ &\le \kappa + \kappa\sigma + \frac{1}{2}L\sigma^2 = \beta. \end{aligned}$$

Hence, from the upper-Lipschitz continuity of $(h+G)^{-1}$, we obtain (18). □

Corollary 3.1 *Let the assumptions of Theorem 3.1 hold, let $v_n = v^*$ for all n, and let σ and κ be the associated constants. Choose σ and κ smaller if necessary so that*

$$\frac{c(\sigma+\kappa)}{(1-c\kappa)} \le 1.$$

Then every Newton sequence $\{x_n\}$ obtained from (16) with $x_0 \in \mathbf{R}^n$ has all elements outside the ball $B_\sigma(x^*)$ or:

(a) x_n is linearly convergent to x^*;

(b) if $H_n \to \nabla f(x^*)$ as $n \to \infty$, then the sequence x_n is superlinearly convergent to x^*;

(c) if $H_n = \nabla f(x_n)$, then x_n is quadratically convergent to x^*.

Proof. Observe that, on our assumptions, if $\{x_n\}$ is a Newton sequence and for some $i \ge 0$ we have $x_i \in B_\sigma(x^*)$, then $x_n \in B_\sigma(x^*)$ for all $n \ge i$. Then apply (18). □

In the remaining part of this section we consider the unperturbed Newton method (5) with a parameter v. Recall that $\mathbf{N}(x, v)$ is the set of all Newton sequences satisfying (5) for v which start from the point x. The following theorem shows the close relation of the map $\mathbf{N}$ with the map $\mathbf{P}$ defined in (9), and hence with the maps $\mathbf{S}$ and $\mathbf{L}$ according to Corollaries 2.1 and 2.2.

Theorem 3.2 *The following are equivalent:*

(i) The map $\mathbf{P}$ *is locally upper-Lipschitz and nonempty-valued at* $((v^*, x^*), x^*)$*;*

(ii) The map $\mathbf{N}$ *is locally upper-Lipschitz and nonempty-valued at* $((x^*, v^*), \xi^*)$.

Proof. (i) $\Rightarrow$ (ii). If $\mathbf{P}$ is locally nonempty-valued, then for each (v, x) close to (v^*, x^*) there exists a Newton step, hence $\mathbf{N}$ is locally nonempty-valued as well. Let $\mathbf{P}$ be locally upper-Lipschitz with constants a and b for the neighborhoods and c and μ for the growth of v and x, respectively. According to Remark 2.1, the constant μ can be made arbitrary small by choosing smaller b. Thus, without loss of generality, suppose that $\mu < 1$. Let $\xi = \{x_1, x_2, \cdots, x_n, \cdots\} \in \mathbf{N}(x, v) \cap B_a(\xi^*)$ for some $(x, v) \in B_b(x^*, v^*)$. Then $x_1 \in \mathbf{P}(v, x) \cap B_a(x^*)$ and from the local upper-Lipschitz continuity of $\mathbf{P}$ we have

$$\|x_1 - x^*\| \leq c\|v - v^*\| + \mu\|x - x^*\|.$$

Proceeding by induction, from $x_{n+1} \in \mathbf{P}(v, x_n) \cap B_a(x^*)$ and from the local upper-Lipschitz continuity of $\mathbf{P}$, we obtain

$$\|x_{n+1} - x^*\| \leq c(1 + \mu + \mu^2 + \cdots + \mu^n)\|v - v^*\| + \mu^{n+1}\|x - x^*\|.$$

Then,

$$\sup_{n \geq 1} \|x_{n+1} - x^*\| \leq \frac{c}{1 - \mu}\|v - v^*\| + \mu\|x - x^*\|.$$

We obtain that the map $\mathbf{N}$ is locally upper-Lipschitz at $((x^*, v^*), \xi^*)$ with constants a, b for the neighborhoods, and growth constants μ for x and $c/(1 - \mu)$ for v.

(ii) $\Rightarrow$ (i). Let $\mathbf{N}$ be locally upper-Lipschitz and locally nonempty-valued at $((x^*, v^*), \xi^*)$ with constants a, b and c. Let $(x, v) \in B_b(x^*, v^*)$ and $z \in \mathbf{P}(v, x) \cap B_a(x^*)$. Then z is the first element of a Newton sequence starting at x for v. From the nonemptiness assumption for $\mathbf{N}$, there exists a Newton sequence $\bar{\xi} = \{\bar{x}_1, \cdots, \bar{x}_n, \cdots\} \in \mathbf{N}(z, v) \cap B_a(\xi^*)$. Observe that the sequence $\eta = \{z, \bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n, \cdots\}$ is an element of $\mathbf{N}(x, v) \cap B_a(\xi^*)$. From the local upper-Lipschitz continuity of $\mathbf{N}$ (in the supremum norm), the first component of η satisfies

$$\|z - x^*\| \leq c(\|x - x^*\| + \|v - v^*\|).$$

This means exactly that $\mathbf{P}$ is locally upper-Lipschitz at $((v^*, x^*), x^*)$. The local nonemptiness of $\mathbf{N}(x, v) \cap B_a(\xi^*)$ implies that $\mathbf{P}$ is locally nonempty-valued. □

Remark 3.1 *Observe that, from the above proof, the local upper-Lipschitz continuity of* **P** *implies the local upper-Lipschitz continuity of* **N** *without the requirement* **P** *be locally nonempty-valued.*

Theorem 3.3 *If the map* **S** *is locally upper-Lipschitz at* (v^*, x^*)*, then* **N** *is locally upper-Lipschitz at* $((x^*, v^*), \xi^*)$*. Conversely, if the map* **N** *is locally upper-Lipschitz and nonempty-valued at* $((x^*, v^*), \xi^*)$*, then* **S** *is locally upper-Lipschitz at* (v^*, x^*)*.*

Proof. If **S** is locally upper-Lipschitz, then, from Corollaries 2.1 and 2.2, the map **P** is locally upper-Lipschitz at $((v^*, x^*), x^*)$; then Theorem 3.2 (with Remark 3.1) completes the proof. Conversely, if **N** is locally upper-Lipschitz and nonempty-valued at $((x^*, v^*), \xi^*)$, then, from Theorem 3.2, **P** is locally upper-Lipschitz at $((v^*, x^*), x^*)$, hence, from Corollaries 2.1 and 2.2, **S** is locally upper-Lipschitz at (v^*, x^*). □

As an illustration, consider the optimization problem (3). The Newton (SQP) iterate x_{n+1} from x_n defined in (6) is a stationary point of the following quadratic program:

$$\min_{z \in C} \left[.5\langle \nabla^2 g(x_n)(z - x_n), z - x_n \rangle + \langle \nabla g(x_n) - v, z - x_n \rangle \right]. \tag{19}$$

Under the second-order sufficient optimality condition, by using the nonemptiness argument in the proof of Theorem 2.1, we conclude that the intersection of the solution set of this problem for (v, x_n) with any ball around x^* is nonempty provided that (v, x_n) is sufficiently close to (v^*, x^*). Thus, the corresponding map **P** is nonempty-valued locally around $((v^*, x^*), x^*)$. The converse follows from Theorems 2.2 and 3.2. Thus we obtain

Corollary 3.2 *The following are equivalent:*

(i) The second-order sufficient optimality condition (10) *holds at* (v^*, x^*)*.*

(ii) The Newton map **N** *defined as in* (19) *is locally upper-Lipschitz and nonempty-valued at* $((x^*, v^*), \xi^*)$ *and* x^* *is a local solution of (3) for* $v = v^*$*.*

Further, from Theorems 2.2, 3.1 and from Corollary 3.2 we obtain that under the second-order sufficient optimality condition, every Newton (SQP) sequence x_n within a sufficiently small neighborhood of the solution x^* is quadratically convergent to x^*; moreover, the map of SQP sequences has the local upper-Lipschitz property. Note that we do not impose any conditions on regularity of the constraints. Also note that, because of the polyhedrality and Robinson's theorem [22], the map from v to the set of all sequences of Lagrange multipliers y_n associated with a Newton (SQP) sequence x_n satisfying (6) has the (standard) upper-Lipschitz property.

For problems with nonlinear smooth constraints, the SQP method is obtained when the Newton method is applied to the variational inequality of the first-order optimality conditions (the Karush-Kuhn-Tucker system). To be specific, consider the problem

$$\min g(z) \quad \text{subject to} \quad \varphi_j(z) \le 0, \quad j = 1, 2, \cdots, m, \quad z \in \mathbf{R}^n, \tag{20}$$

with three times continuously differentiable g and $\varphi = (\varphi_1, \cdots, \varphi_m)$ and with a solution z^* at which the Mangasarian-Fromovitz condition holds (to guarantee the existence of Lagrange multipliers, say y^*). The Karush-Kuhn-Tucker optimality system has the form

$$\begin{aligned} &\nabla g(z) + \nabla \varphi(z)^T y = 0 \\ &\varphi(z) \in N_{\mathbf{R}^n_+}(y). \end{aligned}$$

The SQP method is then the Newton method (5) with $x = (z, y)$,

$$f(x) = \begin{pmatrix} \nabla g(z) + \nabla \varphi^T y \\ \varphi(z) \end{pmatrix} \quad \text{and} \quad F(x) = \begin{pmatrix} 0 \\ N_{\mathbf{R}^n_+}(y) \end{pmatrix}.$$

Note that Corollary 3.1 now claims quadratic convergence of Newton sequences for both the primal variables z and the Lagrange multipliers y. From [8], Theorem 2.6, we obtain that this convergence is guaranteed when the combination of the strict Mangasarian-Fromovitz condition and the standard second-order sufficient optimality condition holds.

As another application of our approach, consider the well-known proximal point method

$$v \in f(x_{n+1}) + \mu(x_{n+1} - x_n) + F(x_{n+1}), \tag{21}$$

where μ is a positive constant. Now the role of the map $\mathbf{P}$ is played by the map $\mathbf{M}_\mu$ defined in (15). Applying Corollary 2.3 we obtain that if the map $\mathbf{S}$ is locally upper-Lipschitz with a growth constant c, then every sequence generated by (21) for v^* and within a sufficiently small ball around x^* is linearly convergent to x^*. Further, let $\mathbf{T}_\mu$ be the map from the pair (x, v) to the set of all sequences obtained by (21) for v and μ and starting from x. The constant sequence $\xi^* = \{x^*, x^*, \cdots\} \in \mathbf{T}_\mu(x^*, v^*)$.

By noting that the growth constant of the map $\mathbf{M}_\mu$ defined in (15) can be < 1 for μ sufficiently small and repeating the argument in Theorem 3.2, we obtain:

Theorem 3.4 *If the mapping* $\mathbf{S}$ *is locally upper-Lipschitz at* (v^*, x^*), *then for a sufficiently small positive* μ *the map* $\mathbf{T}_\mu$ *is locally upper-Lipschitz at* $((x^*, v^*), \xi^*)$.

There is a broad field of applications of our approach to other models in optimization; in particular, to the extended nonlinear programming problem introduced by Rockafellar [24] which covers the conventional nonlinear programming models as well as penalty-function and augmented-Lagrangian type models.

4. AUBIN CONTINUITY

The concept of Aubin continuity can be traced back to the original proofs of the Lyusternik and Graves theorems, see [10, 25] for discussions. In mathematical programming it has appeared as "metric regularity"; in a topological framework, it was called "openness with linear rate". J.-P. Aubin was the first to define this concept as a continuity property of set-valued maps, calling it "pseudo-Lipschitz continuity" [1]. Following [7] we use the name "Aubin continuity".

Definition 4.1 *Let* Γ *map* $\mathbf{R}^n$ *to the subsets of* $\mathbf{R}^m$ *and let* $(y^*, x^*) \in \operatorname{graph}\Gamma$. *We say that* Γ *is Aubin continuous at* (y^*, x^*) *with constants* a, b *for neighborhoods and* c *for growth if for every* $y', y'' \in B_b(y^*)$ *and every* $x' \in \Gamma(y') \cap B_a(x^*)$, *there exists* $x'' \in \Gamma(y'')$ *such that*

$$\|x' - x''\| \le c\|y' - y''\|.$$

Directly from the definition one can extract the following properties of Aubin continuous maps. If a map Γ is Aubin continuous at (y^*, x^*) with constants a, b and c, then for every $0 < a' \le a$ and $0 < b' \le b$ the map Γ is Aubin continuous at (y^*, x^*) with constants a', b' and c. If, in addition, $b' \le a'/c$, then $\Gamma(y) \cap B_{a'}(x^*) \ne \emptyset$ for all $y \in B_{b'}(y^*)$. If Γ is Aubin continuous at (y^*, x^*) with constants a, b and c, there exist constants a', b' and δ such that for every $(y, x) \in \operatorname{graph}\Gamma \cap B_\delta((y^*, x^*))$ the map Γ is Aubin continuous at (y, x) with constants a', b' and c. Finally, for a set-valued map G, if G^{-1} be Aubin continuous at (y^*, x^*) with constants a, b, and c, then for every $\varepsilon < b$ the map

$$y \mapsto \{x \in \mathbf{R}^n \mid y \in G(x) + B_\varepsilon(0)\}$$

is Aubin continuous at (y^*, x^*) with constants a, $b-\varepsilon$, and c. An infinitesimal characterization of the Aubin property is given by Mordukhovich's coderivative criterion, see [25], Chapter 9.

We study the Aubin continuity by following the pattern established in the previous two sections. Our first and basic observation is that, similar to the upper-Lipschitz property, the Aubin continuity is invariant under linearization. This fact is contained in the classical Lyusternik and Graves theorems and their numerous generalizations, see, e.g., [2, 5, 14]. (It is perhaps less known that both Lyusternik and Graves used in their proof a version of the Newton method, the so-called modified Newton method, where, in terms of our notation in (5), $\nabla f(x_n)$ is replaced by $\nabla f(x^*)$. In some of the Russian literature on the subject, see e.g. [3], the modified Newton method is called "Lyusternik process".) This result is presented below in a form very instrumental for applications.

Lemma 4.1 *Let G maps $\mathbf{R}^n$ to the subsets of $\mathbf{R}^m$ and let G^{-1} be Aubin continuous at (v^*, x^*) with constants a and b for neighborhoods and c for growth. Let the nonnegative constants α, β and λ satisfy*

$$\alpha \leq a, \qquad \lambda c < 1, \quad \alpha + \frac{2\beta c}{1 - \lambda c} \leq a, \quad \beta + \lambda\left(\alpha + \frac{2\beta c}{1 - \lambda c}\right) \leq b. \quad (22)$$

Then for every function $h : \mathbf{R}^n \to \mathbf{R}^m$ which is Lipschitz continuous on $B_a(x^*)$ with a Lipschitz constant λ, the map $(h + G)^{-1}$ is Aubin continuous at $(v^* + h(x^*), x^*)$ with constants α and β for neighborhoods and $c/(1 - \lambda c)$ for growth.

Proof. Let $y', y'' \in B_\beta(y^*)$, where $y^* = v^* + h(x^*)$ and let $x' \in (h+G)^{-1}(y') \cap B_\alpha(x^*)$. Then $x' \in G^{-1}(y' - h(x')) \cap B_\alpha(x^*)$ and for both $y = y'$ and $y = y''$ we have

$$\|y - h(x') - v^*)\| \leq \|y - y^*\| + \|h(x') - h(x^*)\| \leq \beta + \lambda\alpha \leq b.$$

From the Aubin continuity of G^{-1} we obtain that there exists $x_2 \in G^{-1}(y'' - h(x'))$ such that

$$\|x_2 - x'\| \leq c\|y' - y''\|.$$

Set $x_1 = x'$. By induction, suppose that there exists a sequence $\{x_k\}$ such that for all $k = 2, 3, \cdots, n$,

$$\|x_k - x_{k-1}\| \leq (\lambda c)^{k-2}\|x_2 - x_1\|$$

and

$$y'' \in h(x_{k-1}) + G(x_k). \qquad (23)$$

Using (22), we have

$$\|x_k - x^*\| \quad \leq \quad \|x_1 - x^*\| + \sum_{j=2}^{k} \|x_j - x_{j-1}\|$$

$$
\begin{aligned}
&\le \|x_1 - x^*\| + \sum_{j=2}^{\infty} \|x_j - x_{j-1}\| \\
&\le \alpha + \frac{c}{1-\lambda c}\|y' - y''\| \\
&\le \alpha + \frac{2\beta c}{1-\lambda c} \le a
\end{aligned} \tag{24}
$$

and

$$\|y'' - h(x_n) - v^*\| \le \|y'' - y^*\| + \|h(x_n) - h(x^*)\| \le \beta + \lambda\left(\alpha + \frac{2\beta c}{1-\lambda c}\right) \le b.$$

Then there exists $x_{n+1} \in G^{-1}(y'' - h(x_n))$ such that

$$
\begin{aligned}
\|x_{n+1} - x_n\| &\le c\|h(x_n) - h(x_{n-1})\| \\
&\le c\lambda (c\lambda)^{n-2}\|x_2 - x_1\| \\
&= (c\lambda)^{n-1}\|x_2 - x_1\|.
\end{aligned}
$$

The induction step is complete. Since $c\lambda < 1$ the sequence $\{x_k\}$ is a Cauchy sequence, hence convergent to, say, x'' which is in $B_a(x^*)$ because of (24). Passing to the limit in (23) we obtain that $x'' \in G^{-1}(y'' - h(x'')) = (h + G)^{-1}(y'')$ and we are done. □

Without changing it, Lemma 4.1 can be stated in abstract spaces, i.e. by replacing $\mathbf{R}^n$ with a complete metric space and $\mathbf{R}^m$ by a linear metric space with a shift-invariant metric.

An application to the maps **S** and **L** defined in Section 2 gives us the following analog to Corollary 2.1:

Corollary 4.1 *The following are equivalent:*

(i) The map **L** *is Aubin continuous at* (v^*, x^*).

(ii) The map **S** *is Aubin continuous at* (v^*, x^*).

The next corollary is a generalization of the previous one; it shows that the Aubin property is persistent when passing from **S** to the map defined by the linearization at a point close to the reference point (v^*, x^*).

Corollary 4.2 *The following are equivalent:*

(i) The map **S** *is Aubin continuous at* (v^*, x^*).

(ii) there exist constants δ, a, b *and* c *such that for every* $(v', x') \in B_\delta((v^*, x^*)) \cap \operatorname{graph} \mathbf{S}$ *the map* $(f(x') + \nabla f(x')(\cdot - x') + F(\cdot))^{-1}$ *is Aubin continuous at* (v', x') *with constants* a, b *and* c.

Proof. Let (i) holds. From the properties of Aubin continuous maps given after Definition 4.1, there exists a $\delta > 0$ such that the map $\mathbf{S}$ is Aubin continuous at any $(v', x') \in B_\delta((v^*, x^*)) \cap \operatorname{graph} \mathbf{S}$ with constants independent of the choice of (v', x'). We now apply Lemma 4.1 with $G = f + F$ and $h(x) = -f(x) + f(x') + \nabla f(x')(x - x')$ at a reference point (v', x'). Let $x', x_1, x_2 \in B_a(x^*)$. Then

$$\|h(x_1) - h(x_2)\| = \| - f(x_1) + f(x_2) - \nabla f(x')(x_1 - x_2)\|$$
$$\leq \int_0^1 \|\nabla f(x_1 + t(x_2 - x_1)) - \nabla f(x')\| dt \|x_1 - x_2\| \leq 2La\|x_1 - x_2\|.$$

Hence, by taking smaller a is necessary, the Lipschitz constant $\lambda = 2La$ of h can be obtained small enough such that $\lambda c < 1$, where c is the growth constant of $\mathbf{S}$. Thus, the map $(h + G)^{-1} = (f(x') + \nabla f(x')(\cdot - x') + F(\cdot))^{-1}$ is Aubin continuous as claimed in (ii).

If (ii) is satisfied, then of course for $(v', x') = (v^*, x^*)$ we obtain that the map $\mathbf{L}$ is Aubin continuous at (v^*, x^*), therefore $\mathbf{S}$ is Aubin continuous at (v^*, x^*), from Corollary 4.1. □

The following is an analog of Corollary 2.2.

Corollary 4.3 *The following are equivalent:*

(i) The map $\mathbf{L}$ *is Aubin continuous at* (v^*, x^*).

(ii) The map $\mathbf{P}$ *is Aubin continuous at* $((v^*, x^*), x^*)$.

Proof. Let (i) hold and let a, b and c be the constants of Aubin continuity of $\mathbf{L}$. Choose α and β such that the inequalities (22) hold with $\lambda = \frac{1}{2}L\beta$. Take α and β smaller if necessary such that

$$L(\beta/4 + 2\alpha) \leq 1. \tag{25}$$

Let (v', x') and (v'', x'') are in $B_{\beta/2}((v^*, x^*))$ and $z' \in \mathbf{P}(x', v') \cap B_\alpha(x^*)$. We apply Lemma 4.1 with

$$G(z) = f(x^*) + \nabla f(x^*)(z - x^*) + F(z)$$

and

$$h(z) = f(x'') + \nabla f(x'')(z - x'') - f(x^*) - \nabla f(x^*)(z - x^*).$$

Let $z_1, z_2 \in B_\alpha(x^*)$. Then

$$\|h(z_1) - h(z_2)\| \leq \|\nabla f(x'') - \nabla f(x^*)\| \|z_1 - z_2\| \leq \frac{1}{2}L\beta\|z_1 - z_2\|,$$

that is, h is Lipschitz continuous with a Lipschitz constant λ. Hence, from Lemma 4.1 the map $(h+G)^{-1}$ is Aubin continuous at $(v^*+h(x^*),x^*)$ with constants α,β and $\gamma=c/(1-\lambda c)$. Note that

$$z' \in (h+G)^{-1}\Big(v'+f(x'')+\nabla f(x'')(z'-x'')-f(x')-\nabla f(x')(z'-x')\Big) \cap B_\alpha(x^*)$$

and

$$\begin{aligned}
&\|v'+f(x'')+\nabla f(x'')(z'-x'')-f(x')-\nabla f(x')(z'-x') \\
&\qquad -v^*-f(x'')-\nabla f(x'')(x^*-x'')+f(x^*)\| \\
\le\ & \|v-v^*\| \\
&+\|f(x^*)-f(x')-\nabla f(x')(x^*-x')+(\nabla f(x')-\nabla f(x''))(x^*-z')\| \\
\le\ & \frac{\beta}{2}+\frac{1}{2}L(\beta/2)^2+L\beta\alpha\le\beta,
\end{aligned}$$

because of (25). Hence, there exists $z''\in(h+G)^{-1}(v'')$ such that

$$\begin{aligned}
&\|z''-z'\| \\
\le\ & \gamma\|v''-v'-f(x'')-\nabla f(x'')(z'-x'')+f(x')+\nabla f(x')(z'-x')\| \\
\le\ & \gamma\|v''-v'\|+\gamma\|f(x'')-f(x')-\nabla f(x')(x''-x')\| \\
&\qquad +\gamma\|(\nabla f(x')-\nabla f(x''))(x''-z')\| \\
\le\ & \gamma\|v''-v'\|+L\gamma(3\beta/2+\alpha)\|x'-x''\|.
\end{aligned}$$

Thus $\mathbf{P}$ is Aubin continuous. The converse implication is immediate by noting that $\mathbf{L}(v)=\mathbf{P}(v,x^*)$. □

Remark 4.1 *By choosing sufficiently small constants α and β for the neighborhoods, the growth constant of $\mathbf{P}$ with respect to x can be made arbitrary small.*

In the remaining part of this section we study the Newton method for Aubin continuous maps. For simplicity, we consider the unperturbed Newton method (5); the analysis can be carried over without major changes to the perturbed form (16). The following theorem shows that the Aubin property implies the existence of Newton sequences for all v close to v^* that are quadratically convergent to a solution $x(v)$ of (1):

Theorem 4.1 *Suppose that the map $\mathbf{S}$ is Aubin continuous at (v^*,x^*). Then there exist constants κ, σ and Θ such that for every $\bar{v}\in B_\kappa(v^*)$*

and every $\bar{x} \in \mathbf{S}(\bar{v}) \cap B_\sigma(x^*)$ the following holds: for every initial point $x_0 \in B_\sigma(x^*)$ there exists a Newton sequence $\{x_n\}$ such that for $n = 0, 1, 2, \cdots$,

$$\|x_{n+1} - \bar{x}\| \leq \Theta \|x_n - \bar{x}\|^2. \tag{26}$$

Proof. Let δ, a, b and c are the constants the existence of which is claimed in Corollary 4.2; that is, for any $(\bar{v}, \bar{x}) \in B_\delta((v^*, x^*)) \cap \operatorname{graph} \mathbf{S}$ the map $(f(\bar{x}) + \nabla f(\bar{x})(\cdot - \bar{x}) + F(\cdot))^{-1}$ is Aubin continuous at $(\bar{v}, \bar{x})$ with constants a, b and c independent of the choice of $(\bar{v}, \bar{x})$. We repeatedly apply Lemma 4.1 with

$$G(x) = f(\bar{x}) + \nabla f(\bar{x})(x - \bar{x}) + F(x)$$

and

$$h(x) = f(x_n) + \nabla f(x_n)(x - x_n) - f(\bar{x}) - \nabla f(\bar{x})(x - \bar{x}).$$

Choose $\sigma = \delta$ and $\kappa = \delta$ and let $\bar{v} \in B_\kappa(x^*)$ and $\bar{x} \in S(\bar{v}) \cap B_\sigma(x^*)$. We will adjust the constant σ after the first step of Newton's method. Let $x_0 \in B_\sigma(x^*)$ and put $n = 0$. Then $h(x) = f(x_0) + \nabla f(x_0)(x - x_0) - f(\bar{x}) - \nabla f(\bar{x})(x - \bar{x})$ and

$$\|h(x_1) - h(x_2)\| \leq \|\nabla f(x_0) - \nabla f(\bar{x})\| \|x_1 - x_2\| \leq 2L\sigma \|x_1 - x_2\|$$

for every $x_1, x_2 \in \mathbf{R}^n$. Note that

$$(\bar{x}, \bar{v} + f(x_0) + \nabla f(x_0)(\bar{x} - x_0) - f(\bar{x})) \in \operatorname{graph}(h + G)$$

and

$$\|f(x_0) + \nabla f(x_0)(\bar{x} - x_0) - f(\bar{x})\| \leq 2L\sigma^2. \tag{27}$$

We apply Lemma 4.1 with the so chosen h (for $n = 0$) and G and with

$$\lambda = 2L\sigma, \quad \alpha = \sigma \quad \text{and} \quad \beta = 2L\sigma^2.$$

It is a matter of simple calculation to adjust the constant σ such that the inequalities (22) hold; for example, we choose σ so small that $2L\sigma c < 1$, etc. Then, from Lemma 4.1, the map $(h + G)^{-1}$ is Aubin continuous at $(\bar{v} + f(x_0) + \nabla f(x_0)(\bar{x} - x_0) - f(\bar{x}), \bar{x})$ with constants α, β and $\gamma = c/(1 - \lambda)$. Hence, from (27) there exists

$$x_1 \in (h + G)^{-1}(\bar{v}),$$

that is, a Newton step from x_0 such that

$$\|x_1 - \bar{x}\| \leq \gamma \| - f(x_0) - \nabla f(x_0)(\bar{x} - x_0) + f(\bar{x})\| \leq 2\gamma L\sigma^2. \tag{28}$$

At this point we take σ smaller if necessary such that

$$\frac{2cL\sigma}{1-2L\sigma} \leq 1.$$

Then, from (28), $x_1 \in B_\sigma(\bar{x})$.

From now on the constants σ is fixed. To obtain a Newton iterate x_2 we apply Lemma 4.1 with the same G and with $h(x) = f(x_1) + \nabla f(x_1)(x - x_1) - f(\bar{x}) - \nabla f(\bar{x})(x - \bar{x})$. Here x_1 will play the role of x_0 in the same way as in the first step obtaining that the map $(h+G)^{-1}$ is Aubin continuous at $(\bar{v}+f(x_1)+\nabla f(x_1)(\bar{x}-x_1)-f(\bar{x}), \bar{x})$ with constants α, β and γ: the same constants as at the first step. Then there exists $x_2 \in (h+G)^{-1}(\bar{v})$, that is, x_2 is a Newton step from x_1 for $\bar{v}$, such that

$$\|x_2 - \bar{x}\| \leq \gamma\| - f(x_1) - \nabla f(x_1)(\bar{x} - x_1) + f(\bar{x})\| \leq \frac{1}{2}\gamma L\|x_1 - \bar{x}\|^2.$$

The Newton iterates x^3, x^4, $\cdots$ are obtained in the same way. Taking $\Theta = \gamma L/2$ the proof is complete. □

Recall that $\mathbf{N}(x, v)$ is the set of all Newton sequences, that is, sequences that satisfy (5) starting from the point x and associated with the value v of the parameter. The constant sequence ξ^* is an element of $\mathbf{N}(x^*, v^*)$.

Theorem 4.2 *Suppose that the map* $\mathbf{S}$ *is Aubin continuous at* (v^*, x^*). *Then the map* $\mathbf{N}$ *is Aubin continuous at* $((x^*, v^*), \xi^*)$.

Proof. From Corollaries 4.1 and 4.3, the map $\mathbf{P}$ is Aubin continuous at $((v^*, x^*), x^*)$; let a, b be the associated constants for neighborhoods, c the growth constant with respect to v and μ the growth constant with respect to x. By taking a and b smaller if necessary, we can have $\mu < 1$. Let $v', v'' \in B_b(v^*)$ and $x', x' \in B_b(x^*)$, and let $\xi' \in \mathbf{N}(x', v') \cap B_a(\xi^*)$, $\xi' = \{x_1', x_2', \cdots, x_n', \cdots\}$. Note that $x_1' \in \mathbf{P}(v', x') \cap B_a(x^*)$. Then there exists $x_1'' \in \mathbf{P}(v'', x'')$ such that

$$\|x_1'' - x_1'\|| \leq c\|v'' - v'\| + \mu\|x'' - x'\|.$$

Further, $x_2' \in \mathbf{P}(v_1', x') \cap B_a(x^*)$, hence there exists $x_2'' \in \mathbf{P}(v'', x_1'')$ such that

$$\|x_2'' - x_2'\|| \leq c\|v'' - v'\| + \mu\|x_1'' - x_1''\| \leq c(1+\mu)\|v'' - v'\| + \mu^2\|x'' - x'\|.$$

By induction, we obtain that there exists a sequence $\xi'' \in \mathbf{N}(x'', v'')$, $\xi'' = \{x_1'', x_2'', \cdots, x_n'', \cdots\}$. such that

$$\|x_n'' - x_n'\|| \leq c(1 + \mu + \mu^2 + \cdots + \mu^{n-1})\|v'' - v'\| + \mu^n\|x'' - x'\|.$$

Taking the supremum norm we obtain that $\mathbf{N}$ is Aubin continuous with constants a and b for the neighborhoods and growth constants μ for x and $c/(1-\mu)$ for v. □

In the above theorem $\mathbf{N}(x,v)$ is the set of *all* Newton sequences starting from x for v. From Theorem 4.1 we know that at least one of these sequences will be convergent provided that x is sufficiently close to x^* and v is sufficiently close to v^*. Denote by $\mathcal{N}(x,v)$ the subset of $\mathbf{N}(x,v)$ containing all *convergent* Newton sequences starting from the point x and associated with v. The following theorem is a kind of converse to Theorem 4.2.

Theorem 4.3 *If the map $\mathcal{N}$ is Aubin continuous at $((x^*,v^*),\xi^*)$, then the map $\mathbf{S}$ is Aubin continuous at (v^*,x^*).*

Proof. Let $\mathcal{N}$ be Aubin continuous with constants a, b and c. Let $v',v'' \in B_b(v^*)$ and let x' be an element of $\mathbf{S}(v') \cap B_a(x^*)$. The constant sequence $\xi' = \{x',x',\cdots,x',\cdots\}$ is an element of $\mathcal{N}(x',v') \cap B_a(\xi^*)$. From the Aubin property of $\mathcal{N}$ there exists $\xi'' = \{x''_1,x''_2,\cdots,x''_n,\cdots\} \in \mathcal{N}(x',v'')$, such that for every $n \geq 1$

$$\|x''_n - x'_n\| \leq c\|v''-v'\|. \tag{29}$$

Since $\xi'' \in \mathcal{N}(x',v'')$, ξ'' is convergent, say to x''. Clearly, every convergent Newton sequence for v'' is convergent to a solution of (1) for v''. Then x'' is an element of $\mathbf{S}(v'')$. Passing to the limit in (29) we obtain that

$$\|x''-x'\| \leq c\|v''-v'\|.$$

Thus $\mathbf{S}$ is Aubin continuous. □

Let us apply the above result to our illustrative optimization problem (3). The basic condition here will be the Aubin continuity of the map $\mathcal{X}$ defined in (11) (or, equivalently, of the map $\mathcal{L}$ defined in (12)). If we assume the Aubin property for $\mathcal{L}$, however, we end up having $\mathcal{L}$ locally singe-valued because of the following result obtained in [7]: *Let A be an $n \times n$ real matrix and let C be a convex polyhedral set. Then the map $(A+N_C)^{-1}$ is Aubin continuous at (v^*,x^*) if and only if it is locally single-valued and Lipschitz continuous around (v^*,x^*).*

Using the terminology of [25], we call the property just described Lipschitzian localization; we consider it in the following section.

5. LIPSCHITZIAN LOCALIZATION

The Lipschitzian localization property simply says that when restricted to a neighborhood of a point in its graph, a possibly set-valued map becomes a Lipschitz continuous (single-valued) function.

Definition 5.1 *Let Γ maps $\mathbf{R}^n$ to the subsets of $\mathbf{R}^m$ and let $(y^*, x^*) \in$ graph Γ. We say that Γ has a Lipschitzian localization at (y^*, x^*) with constants a, b and c if the map $y \mapsto \Gamma(y) \cap B_a(x^*)$ is single-valued (a function) and Lipschitz continuous in $B_b(y^*)$ with a Lipschitz constant c.*

If a map Γ has a Lipschitzian localization at (y^*, x^*), then it is Aubin continuous at (y^*, x^*) with the same constants. Conversely, if Γ is Aubin continuous at (y^*, x^*) with constants a, b and c and, in addition, for some positive constants α and β, $\Gamma(y) \cap B_\alpha(x^*)$ consists of at most one point for every $y \in B_\beta(y^*)$, then Γ has a Lipschitzian localization at (y^*, x^*) with constants a', b', c' provided that

$$0 < a' < \min\{a, \alpha\} \quad \text{and} \quad 0 < b' \leq \min\{b, \beta, a'/c', (\alpha - a')/c'\}.$$

The following lemma is an inverse function theorem which can be proved in various ways. Its statement is parallel to Lemmas 2.1 and 4.1.

Lemma 5.1 *Let G maps $\mathbf{R}^n$ to the subsets of $\mathbf{R}^m$ and let G^{-1} has a Lipschitzian localization at (v^*, x^*) with constants a and b for neighborhoods and c for growth. Let the nonnegative constants α, β and λ satisfy*

$$\alpha \leq a, \qquad \lambda c < 1, \quad c(\beta + \lambda\alpha) \leq \alpha, \quad \beta + \lambda\alpha \leq b. \tag{30}$$

Then for every function $h : \mathbf{R}^n \to \mathbf{R}^m$ which is Lipschitz continuous on $B_\alpha(x^*)$ with a Lipschitz constant λ, the map $(h+G)^{-1}$ has a Lipschitzian localization at $(v^* + h(x^*), x^*)$ with constants α and β for neighborhoods and $c/(1 - \lambda c)$ for growth.

Sketch of Proof. Let α, β, λ satisfy (30) and let h be Lipschitz continuous on $B_a(x^*)$ with a Lipschitz constant λ. Let $v \in B_\beta(v^*)$ and let $y = v + h(x^*)$, $y^* = v^* + h(x^*)$. Then $\|y - h(x) - y^*\| \leq \|v - v^*\| + \|h(x) - h(x^*)\| \leq \beta + \lambda\alpha \leq b$. It is easy to see that, for each $v \in B_\beta(v^*)$, the map

$$x \mapsto \Phi(x) := G^{-1}(y - h(x)) \cap B_\alpha(x^*)$$

is a function from $B_a(x^*)$ to $B_a(x^*)$ and is a contraction, hence it has a unique fixed point, say, $x(y)$, in $B_\alpha(x^*)$. Since $x(y) \in (G + h)^{-1}(y) \cap$

$B_\alpha(x^*)$, the map $(G+h)^{-1} \cap B_\alpha(x^*)$ is a function defined (with nonempty values) on $B_\beta(y^*)$. Moreover, for any $y', y'' \in B_\beta(y^*)$ we have

$$\begin{aligned} &\|x(y') - x(y'')\| \\ = \;& \|G^{-1}(y' - h(x(y'))) \cap B_\alpha(x^*) - G^{-1}(y'' - h(x(y''))) \cap B_\alpha(x^*)\| \\ \leq \;& c(\|y' - y''\| + \lambda\|x(y') - x(y'')\|), \end{aligned}$$

hence

$$\|x(y') - x(y'')\| \leq \frac{c}{1 - \lambda c}\|y' - y''\|.$$

Thus $y \mapsto x(y)$ is Lipschitz continuous with a Lipschitz constants $c/(1 - \lambda c)$. □

Applying Lemma 5.1 to the maps **S**, **L** and **P** defined in Section 2 we obtain:

Corollary 5.1 *The following are equivalent:*

(i) The map **L** *has a Lipschitzian localization at* (v^*, x^*).

(ii) The map **S** *has a Lipschitzian localization at* (v^*, x^*)

(iii) The map **P** *has a Lipschitzian localization at* $((v^*, x^*), x^*)$.

In optimization, the existence of Lipschitzian localization of the linearization map **L** is often called "strong regularity", a concept introduced by Robinson; actually, the implication (i) ⇒ (ii) is essentially the contents of Robinson's implicit function theorem in [21]. A characterization of the Lipschitzian localization property of a locally continuous set-valued map in finite dimensions is provided by the so-called strict graphical derivative, see [25], Theorem 9.54. A far reaching recent study of optimization problems with solution maps having this property is given in [18].

For our illustrative problem (3) the Lipschitzian localization property of the solution map is equivalent to the so-called *strong second-order sufficient optimality condition* which has the form:

$$\langle u, \nabla^2 g(x^*)u\rangle > 0 \quad \text{for all nonzero} \quad u \in K - K, \tag{31}$$

where K the critical cone at (v^*, x^*) defined in Section 2. Specifically, we have the following result which can be extracted from [7] or from more recent papers [13, 17, 18]:

Theorem 5.1 *The following are equivalent for the problem (3):*

(i) *The strong second-order sufficient optimality condition (31) holds at* (v^*, x^*).

(ii) *The Argmin map has a Lipschitzian localization at* (v^*, x^*).

By combining Theorem 5.1 with the last property of Aubin continuous maps given after Definition 4.1 and the result in [7] given at the end of Section 4, we obtain that the strong second-order sufficient optimality condition is equivalent to the property that for every sufficiently small $\varepsilon > 0$ the ε-stationary map

$$v \mapsto \mathcal{X}_\varepsilon(v) = \{x \in \mathbf{R}^n \mid \text{dist}\,(v - \nabla g(x), N_C(x)) \leq \varepsilon\}$$

is Aubin continuous at (v^*, x^*).

Let us consider the Newton method (5) applied to the general model (1). The analog of Corollary 3.1 and Theorem 4.1 has the following form:

Theorem 5.2 *Suppose that the map* $\mathbf{S}$ *has a Lipschitzian localization at* (v^*, x^*). *Then there exist constants* κ, σ *and* Θ *such that for every* $\bar{v} \in B_\kappa(v^*)$ *and every initial point* $x_0 \in B_\sigma(x^*)$ *there exists a unique Newton sequence* $\{x_n\}$ *and this sequence is quadratically convergent to* $\bar{x} := \mathbf{S}(\bar{v}) \cap B_\sigma(x^*)$ *with a constant* Θ, *that is,*

$$\|x_{n+1} - \bar{x}\| \leq \Theta \|x_n - \bar{x}\|^2, \quad n = 0, 1, 2, \cdots.$$

Proof. The proof is completely analogous to the proofs of Theorem 4.1 with Lemma 5.1 replacing Lemma 4.1. □

The next result is the analog of Theorems 3.3, 4.3 and 4.4.

Theorem 5.3 *The following are equivalent:*

(i) *The map* $\mathbf{S}$ *has a Lipschitzian localization at* (v^*, x^*).

(ii) *The Newton map* $\mathbf{N}$ *has a Lipschitzian localization at* $((x^*, v^*), \xi^*)$.

Applied to the illustrative optimization problem (3) and the SQP method (6), we obtain that the strong second-order sufficient optimality condition (31) at (v^*, x^*) is equivalent to the following: for every v near v^* and x near x^* there exists a unique sequence of primal variables x_n starting from x and which is quadratically convergent to the unique solution $x(v)$ of (3) for v, for some sequence of Lagrange multipliers y_n. Moreover the map from (x, v) to the set of all sequences x_n staring from x for v has a Lipschitzian localization at $((x^*, v^*), \xi^*)$.

For the nonlinear program (20), the Newton (SQP) step is performed for the pair (x, y) with x being the primal variable and y the Lagrange multiplier. Theorem 5.4 extends the characterization of the Lipschitzian localization property of the Karush-Kuhn-Tucker map to the SQP map. Specifically, if we assume that the reference solution is isolated, then the Lipschitzian localization property of the SQP method is equivalent to the combination of the linear independence condition for the active constraints and the strong second-order sufficient optimality condition at the reference point.

References

[1] J.-P. Aubin (1984), Lipschitz behavior of solutions to convex minimization problems, *Math. Oper. Res.*, 9, pp. 87–111.

[2] R. Cominetti (1990), Metric regularity, tangent sets, and second-order optimality conditions, *Appl. Math. Optim.*, 21, pp. 265–287.

[3] A.V. Dmitruk, A.A. Milyutin and N.P. Osmolovskiĭ (1980), The Lyusternik theorem and the theory of extremum, *Uspekhi Math. Nauk*, 35, pp. 11–46.

[4] A.L. Dontchev (1995), Characterizations of Lipschitz stability in optimization, in *Recent developments in well-posed variational problems,*, Math. Appl., 331, Kluwer, Dordrecht, pp. 95–115.

[5] A.L. Dontchev (1996), The Graves theorem revisited, *J. Convex Anal.*, 3, pp. 45–53.

[6] A.L. Dontchev, W.W. Hager and V. Veliov, Uniform convergence and mesh independence of Newton's method for discretized variational problems, *SIAM J. Control and Optim.,* to appear.

[7] A.L. Dontchev and R.T. Rockafellar (1996), Characterizations of strong regularity for variational inequalities over polyhedral convex sets, *SIAM J. Optim.*, 6, pp. 1087–1105.

[8] A.L. Dontchev and R.T. Rockafellar (1998), Characterizations of Lipschitzian stability in nonlinear programming, in *Mathematical programming with data perturbations,* Lecture Notes in Pure and Appl. Math., 195, Dekker, New York, pp. 65–82.

[9] W.W. Hager and M.S. Gowda (1999), Stability in the presence of degeneracy and error estimation, *Math. Programming*, 85, pp. 181–192.

[10] A.D. Ioffe, Metric regularity and subdifferential calculus, to appear.

[11] A.J. King and R.T. Rockafellar (1992), Sensitivity analysis for nonsmooth generalized equations, *Math. Programming*, 5, Ser. A, pp. 193–212.

[12] D. Klatte, Upper Lipschitz behavior of solutions to perturbed $C^{1,1}$ programs, preprint.

[13] D. Klatte and B. Kummer (1999), Strong stability in nonlinear programming revisited, *J. Austral. Math. Soc.*, 40, Ser. B, pp. 336–352.

[14] B. Kummer (1998), Lipschitzian and pseudo-Lipschitzian inverse functions and applications to nonlinear optimization, in *Mathematical programming with data perturbations,* Lecture Notes in Pure and Appl. Math., 195, Dekker, New York, pp. 201–222.

[15] A.B. Levy (1996), Implicit multifunction theorems for the sensitivity analysis of variational conditions, *Math. Programming*, 74, Ser. A, pp. 333–350.

[16] A. Levy, Solution sensitivity from general principles, preprint.

[17] R.A. Poliquin and R.T. Rockafellar (1998), Tilt stability of local minima, *SIAM J. Optim.*, 2, pp. 287–299.

[18] A. Levy, R.A. Poliquin and R.T. Rockafellar, Stability of locally optimal solutions, *SIAM J. Optim.,* to appear.

[19] J. Renegar (1995), Linear programming, complexity theory and elementary functional analysis, *Math. Programming*, 70, Ser. A, pp. 279–351.

[20] S.M. Robinson (1979), Generalized equations and their solution, Part I: Basic Theory, *Math. Programming Study*, 10, pp. 128–141.

[21] S.M. Robinson (1980), Strongly regular generalized equations, *Math. Oper. Res.*, 5, pp. 43–62.

[22] S.M. Robinson (1981), Some continuity properties of polyhedral multifunctions, *Math. Programming Study*, 14, pp. 206–214.

[23] R.T. Rockafellar (1989), Proto-differentiability of set-valued mappings and its applications in optimization, *Analyse non linéaire (Perpignan, 1987), Ann. Inst. H. Poincaré Anal. Non Linéaire* 6, suppl., pp. 449–482.

[24] R.T. Rockafellar, Extended nonlinear programming, in *Nonlinear Optimization and Applications 2*, Kluwer, Dordrecht, to appear.

[25] R.T. Rockafellar and R.J.-B. Wets (1997), Variational analysis, Springer-Verlag, Berlin.

[26] H. Xu (1999), Set-valued approximations and Newton's methods, *Math. Programming*, 84, Ser. A, pp. 401–420.

[27] T. Zolezzi (1998), Well-posedness and conditioning of optimization problems, *Proc. 4th International Conf. Math. Methods Oper. Research and 6th Workshop on Well-posedness and Stability of Optimization Problems, Pliska Stud. Math. Bulgar.*, 12, pp. 267–280.