ON THE DOMAIN OF THE IMPLICIT FUNCTION AND APPLICATIONS

MARCO PAPI

Received 31 July 2003 and in revised form 22 December 2004

The implicit function theorem asserts that there exists a ball of nonzero radius within which one can express a certain subset of variables, in a system of equations, as functions of the remaining variables. We derive a lower bound for the radius of this ball in the case of Lipschitz maps. Under a sign-preserving condition, we prove that an implicit function exists in the case of a set of inequalities. Also in this case, we state an estimate for the size of the domain. An application to the local Lipschitz behavior of solution maps is discussed.

1. Introduction

The implicit function theorem is one of the fundamental results in multivariable analysis [1, 8, 11]. It asserts that if $F_i(x,y)$, $i = 1,\dots,n$, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, are countinuously differentiable functions in a neighborhood of a point (x_0, y_0), where $F_i(x_0, y_0) = 0$, for $i = 1,\dots,n$, and the Jacobian

$$D_yF(x_0, y_0) = \left(\frac{\partial F_i}{\partial y_j}(x_0, y_0)\right)_{1\le i,j\le n} \tag{1.1}$$

is invertible, then there exist a positive number $r > 0$ and continuous functions $g_1(x),\dots, g_n(x)$, defined in the domain $B = \{x \in \mathbb{R}^m : |x - x_0| < r\}$, such that $g_i(x_0) = y_{0i}$ and $F_i(x,g_1(x),\dots,g_n(x)) = 0$, for $i = 1,\dots,n$, in B. This theorem has been extended to Lipschitz functions by Clarke [4, 5]. In this case $F = (F_1,\dots,F_n)$ is a locally Lipschitz function in a neighborhood of (x_0, y_0) and the invertibility assumption is required for all the matrices of the generalized Jacobian of F at (x_0, y_0).

Despite the central role played by this result in analysis, multidimensional nonlinear optimization algorithms [2, 7, 16, 17], and in developing Newton-type methods for solving nonsmooth equations [12, 13, 18], a lower bound for the size of the domain B has not been sufficiently investigated in the literature. The first nontrivial estimate has been reported in [3] for the case of complex analytic functions. The authors base their result on the Roche theorem to derive a lower bound in the case $n = 1$, then they recursively extend this estimate to the general case.

Journal of Inequalities and Applications 2005:3 (2005) 221–234
DOI: 10.1155/JIA.2005.221

Some important problems, like those which appear in sensitivity and stability analysis of systems of equations and inequalities [14, 15], do not show strong regularity properties. Therefore, over the years, a great deal of attention has been focused on developing new tools for maps not necessarily differentiable.

Given the high relevance played by the Lipschitz continuity, one of the purposes of this paper is to establish an estimate for the size of the domain B and consequently of the set of values of the implicit function, in the context of Lipschitz continuous maps. These estimates can be applied for proving the upper-Lipschitz continuity [19] of some set-valued maps. This is a recently introduced concept of regularity which turns out to be quite natural in nonlinear optimization. For illustration, we consider the question of the local Lipschitz behavior of the map "parameter x maps to set of solutions of $x \in f(y)+C$," where $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}^m$ is a Lipschitz function and $C \subset \mathbb{R}^m$ is closed; more precisely, we consider the following map:

$$x \in \mathbb{R}^m \longmapsto \mathbf{S}(x) = \{y \in \mathbb{R}^n : x \in f(y)+C\}. \tag{1.2}$$

In particular, when $C = \{0\}$, we obtain a system of equations. For $C = \mathbb{R}^m_+$, the positive orthant in $\mathbb{R}^m$, we have a system of inequalities. The local Lipschitz properties of $\mathbf{S}$, are mainly used for constructing effective numerical algorithms. Actually, by Newton's method [18] applied to the problem $x \in f(y)+C$, we mean the following procedure which generates a sequence $\{y_1, y_2, \ldots\}$, with a given starting point y_0, according to the rule

$$x \in f(y_n)+\partial f(y_n)\cdot(y_{n+1}-y_n)+C, \tag{1.3}$$

where ∂f denotes the generalized Jacobian of f. For given x and y, let $\{y_n\}$ be a Newton sequence, that is, a sequence starting from y and satisfying (1.3). Denote by $\mathbf{N}(x,y)$ the set of all Newton sequences for x, starting from the point y. Then in [6] it is proved that the local upper-Lipschitz continuity of $\mathbf{S}$ implies that every Newton sequence, within a sufficiently small ball around the solution, is convergent. Moreover the radius of this ball is controlled by a constant which depends on the local Lipschitz behavior of $\mathbf{S}$. Actually, in establishing the upper-Lipschitz continuity, we have to estimate three parameters which characterize this property and, as shown in [6], these numbers can be used to derive the rate of convergence of Newton sequences. As shown in Section 4, the implicit function theorems we prove can be used to find lower bounds for these three constants, see Proposition 4.2.

In many applications, as in the problem mentioned above, we are mainly interested in finding an implicit function for a set of inequalities (i.e., $F_i \le 0$, for $1 \le i \le n$), where the variable y is constrained to stay in a closed convex set $\Omega \subset \mathbb{R}^n$. In this case, we cannot apply the classical version of the implicit function theorem because the implicit function has to map the variable x into the set Ω. Moreover the reference point y_0 can lie on the boundary of Ω, while the map F could not be continuously differentiable in a neighborhood of this point, as required by the classical version of the implicit function theorem. Another interesting case appears when the map F is Lipschitz continuous around the reference point (x_0, y_0), but the generalized Jacobian at this point contains singular matrices. In such a situation, we cannot apply the Lipschitz version of that result, see Example 4.7.

Then the question is can we construct an implicit function $g(x) = (g_1(x),\dots,g_n(x))$ such that $F_i(x,g(x)) \le 0$, for $1 \le i \le n$, in some neighborhood of x_0?

A second purpose of this work, is to give an answer to this question. By requiring a sign-preserving condition on the Jacobian, we will prove that an implicit function exists, see Theorem 3.4. This result can be used to study the local Lipschitz properties of the solution map (1.2). Therefore, also for this version of the implicit function theorem, we state a lower bound for the size of the domain of the implicit function.

The outline of the paper is as follows. Section 2 provides some basic definitions and notations used in the rest of the paper. In Section 3, we state our main results. In Section 4, we discuss the connection between our extensions of the implicit function theorem and the local Lipschitz behavior of the solution map (1.2). Furthermore, for an easier comprehension of our assumptions and results, in this section, we present some examples. Finally Section 5 is devoted to the proof of the results.

2. Main notations and definitions

In this section, we introduce the main notations and definitions used in this paper.

(1) Given a separable metric space (M,d), where d is a distance on M, the open ball in M with center $x \in M$ and radius $r > 0$ is denoted by $B_r(x) = \{y \in M : d(y,x) < r\}$. If $M = \mathbb{R}^m$, we take for d the norm $|\cdot|$, obtained by the usual scalar product $\langle\cdot,\cdot\rangle$ on $\mathbb{R}^m$. Moreover, for any set $A \subset M$, $\operatorname{cl}A$, ∂A denote the closure and the boundary of A, respectively; for every $\varepsilon > 0$, $A^\varepsilon = \{y : d(y,A) < \varepsilon\}$ denotes the open ball of radius ε around A.

(2) Given an open set $A \subset \mathbb{R}^m$ and $B \subset \mathbb{R}^n$, the space $C^k(A;B)$, for $k \ge 0$, denotes the space of B-valued functions that are continuous with their derivatives up to the order k in A. The space $C^k(\operatorname{cl}A;B)$ is the space of functions in $C^k(A;B)$, such that each derivative up to the order k can be continuously extended to ∂A. When $B = \mathbb{R}^n$, we will omit to indicate the set of values. We will use the notations D and D_x for the Jacobians of a vector-valued function, where the subscript stands for the partial derivation with respect to the variable x.

Let $F \in C^1(\operatorname{cl}A)$ and $x \in \partial A$; in the sequel, $DF(x)$ will denote the value at x of the continuous extension, to the closure of A, of the usual Jacobian of F.

(3) Let $x \in \mathbb{R}^n$, then we say $x \le 0$ (resp., $x \ge 0$) if and only if $x_i \le 0$ (resp., $x_i \ge 0$) for every $i = 1,\dots,n$.

(4) Let $\mathbb{R}^{m\times n}$ be the space of all real matrices with m rows and n columns. We will denote by I_m the identity matrix of order m. Let $S \in \mathbb{R}^{m\times n}$, then we define the norm of S as

$$\|S\| = \max_{\substack{x\in\mathbb{R}^n \\ |x|=1}} |S\cdot x|. \tag{2.1}$$

Let $\mathcal{S}$ be a set of real square matrices of the same order. We say that $\mathcal{S}$ is invertible if every matrix in $\mathcal{S}$ is invertible and we will denote by $\mathcal{S}^{-1}$ the inverse of $\mathcal{S}$, that is,

$$\mathcal{S}^{-1} = \{S^{-1} : S \in \mathcal{S}\}. \tag{2.2}$$

We extend the norm (2.1) to the case of a bounded set of matrices $\mathcal{S}$:

$$\|\mathcal{S}\|_\infty = \sup_{S\in\mathcal{S}} \|S\|. \tag{2.3}$$

For every n and $\lambda > 0$, let

$$\mathcal{S}(\lambda, n) = \{S \in \mathbb{R}^{n\times n} : S \text{ is invertible and } \|S^{-1}\| < \lambda\}. \tag{2.4}$$

Definition 2.1. Let $S \in \mathbb{R}^{n\times n}$, then S is *sign preserving* if and only if the following holds true:

$$S \cdot p \geq 0 \quad \forall p \in \mathbb{R}^n \text{ s.t. } p \geq 0. \tag{2.5}$$

We recall the notion of generalized Jacobian for Lipschitz functions introduced by F. H. Clarke.

(5) Let F be an $\mathbb{R}^m$-valued Lipschitz continuous function over an open domain $A \subset \mathbb{R}^n$. We define the Lipschitz constant of F as

$$\operatorname{Lip}(F) = \sup_{\substack{x,x'\in A\\ x\neq x'}} \frac{|F(x) - F(x')|}{|x - x'|}. \tag{2.6}$$

Let $D(F)$ denote the set of all points in A where F is differentiable, then by Rademacher's theorem, $D(F)$ is of full Lebesgue measure on A. Therefore we can define the generalized Jacobian of F at a point $x \in A$:

$$\partial F(x) = \operatorname{co}\Big\{P \in \mathbb{R}^{m\times n} : P = \lim_{x_h\to x} DF(x_h),\ x_h \in D(F)\Big\}, \tag{2.7}$$

where the notation "co" indicates the convex hull. We define the generalized partial derivative of F at $x = (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, with $n_1 + n_2 = n$, as

$$\partial_{x_1} F(x_1, x_2) = \{P \in \mathbb{R}^{m\times n_1} : \exists Q \in \mathbb{R}^{m\times n_2}, \text{ s.t. } [P, Q] \in \partial F(x)\}. \tag{2.8}$$

In a similar way, one can define $\partial_{x_2} F(x_1, x_2)$. Some fundamental properties of the generalized Jacobian are summarized below.

(6) Let Γ be a set-valued map from $A \subset \mathbb{R}^m$ to the subsets of $\mathbb{R}^n$. For every $U \subset A$, we set

$$\Gamma(U) = \bigcup_{x\in U} \Gamma(x). \tag{2.9}$$

Definition 2.2. Let Γ be a set-valued map from $A \subset \mathbb{R}^m$ to the subsets of $\mathbb{R}^n$. Γ is said to be upper semicontinuous (u.s.c.) at $x \in A$ if, for every neighborhood V of $\Gamma(x)$ in $\mathbb{R}^n$, there exists a neighborhood U of x in $\mathbb{R}^m$ such that $\Gamma(U \cap A) \subset V$.

Proposition 2.3 [5, Proposition 2.6.2]. *Let $F : A \subset \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz continuous near $x \in A$. Then the following statements hold:*

(a) $\partial F(x)$ *is a nonempty convex compact subset of* $\mathbb{R}^{m\times n}$;

(b) $\partial F(x)$ *is closed at* x;

(c) $\partial F(x)$ *is u.s.c. at* x*. In particular, for any* $\varepsilon > 0$*, there exists* $\delta > 0$ *such that*

$$\partial F(x') \subset \partial F(x) + \varepsilon B_1(0) \quad \forall x' \in B_\delta(x) \cap A. \tag{2.10}$$

We recall a key result in the theory of Lipschitz continuous functions.

Theorem 2.4 (mean value theorem, [5, Proposition 2.6.5]). *Let $F : A \to \mathbb{R}^m$ be Lipschitz continuous on an open convex set $A \subset \mathbb{R}^n$, and let $x, y \in A$. Then, there exists a matrix $P \in \mathrm{co}\partial F([x, y])$ (where $[x, y]$ stands for the straight-line segment connecting x and y) such that*

$$F(x) - F(y) = P \cdot (x - y). \tag{2.11}$$

3. Main results

In this section, we present our main results. The first version of the implicit functions theorem is concerned with the case of Lipschitz maps, and here we estimate the size of the neighborhoods where the implicit function is defined. The proof of this result is based on a fundamental inclusion proved in Proposition 5.1.

As explained in Section 4, the other result we present allows to deal with the following situations.

(1) The reference point y_0 lies on the boundary of the set where the map F is defined.

(2) The map F is Lipschitz continuous around the reference point but the generalized Jacobian is not invertible (see point (4) in Section 2).

In these cases, assuming a sign-preserving condition (see Definition 2.1), we prove the existence of an implicit function for the system of inequalities $F \leq 0$. Furthermore, as for the previous result, we state a lower bound for the size of the domain where this function is defined.

Theorem 3.1. *Let $F : \mathbb{O} \to \mathbb{R}^n$ be a Lipschitz map defined in the open set $\mathbb{O} \subset \mathbb{R}^m \times \mathbb{R}^n$. Let $(x_0, y_0) \in \mathbb{O}$, $L > 0$, $\lambda > 0$, and $r > 0$ be such that $B_r(x_0) \times B_r(y_0) \subset \mathbb{O}$ and the following hypotheses hold:*

(i) $F(x_0, y_0) = 0$;

(ii) $\|(\partial_y F(x_0, y_0))^{-1}\|_\infty < \lambda$;

(iii) $\|\partial_x F(B_r(x_0) \times B_r(y_0))\|_\infty \leq L$.

Let $r_2 = r_1/2\ell$, $\ell = \max(1, (1 + L)\lambda)$, where

$$r_1 = \sup\{\rho \in [0, r] : \mathrm{co}\{\partial_y F(B_\rho(x_0) \times B_\rho(y_0))\} \subset \mathcal{S}(\lambda, n)\}. \tag{3.1}$$

Then there exists a unique function $g \in C(B_{r_2}(x_0); B_{r_1}(y_0))$ satisfying $g(x_0) = y_0$ and

$$F(x, g(x)) = 0 \quad \forall x \in B_{r_2}(x_0). \tag{3.2}$$

Furthermore the following inequality holds true:

$$|g(x_1) - g(x_2)| \leq \lambda L |x_1 - x_2| \quad \forall x_1, x_2 \in B_{r_2}(x_0). \tag{3.3}$$

Remark 3.2. By the properties recalled in Proposition 2.3, it is clear that the set-valued map $\rho \mapsto \text{co}\{\partial_y F(B_\rho(x_0) \times B_\rho(y_0))\}$, defined for $\rho \in [0,r]$, is u.s.c. at 0. By (ii), $\partial_y F(x_0, y_0) \subset \mathcal{S}(\lambda, n)$ and by (2.4), $\mathcal{S}(\lambda, n)$ is open in $\mathbb{R}^{n\times n}$, therefore the number r_1 in (3.1) is well defined. We observe that the radius r_2 represents a lower bound for the size of the domain of the implicit function.

Remark 3.3. The uniqueness of the implicit function g in Theorem 3.1 implies that any point $(x, y) \in B_{r_2}(x_0) \times B_{r_1}(y_0)$ such that $F(x, y) = 0$ belongs to the graph of g, that is, $g(x) = y$.

Theorem 3.4. *Let $F : \mathbb{O} \to \mathbb{R}^n$ be a continuous function in the open set $\mathbb{O} \subset \mathbb{R}^m \times \mathbb{R}^n$. Let $\varepsilon > 0$, $(x_0, y_0) \in \mathbb{O}$, r_1, $r_2 > 0$ be such that, if $B_1^+ = \{y \in \mathbb{R}^n : 0 \le y - y_0 \le r_1\}$ and $B_2 = \{x \in \mathbb{R}^m : |x - x_0| \le r_2\}$, then $B_2 \times B_1^+ \subset \mathbb{O}$ and the following conditions hold:*

($\jmath$) *$F(x_0, y_0) \le 0$;*

($\jmath\jmath$) *$F \in C^1(B_2 \times B_1^+)$;*

($\jmath\jmath\jmath$) *$-D_y F(x_0, y_0)$ is invertible and sign preserving;*

($\jmath\jmath\jmath\jmath$) *the following inequalities are satisfied:*

$$\sup_{x\in B_2} |F(x, y_0) - F(x_0, y_0)| \le \frac{r_1}{(1+\varepsilon\sqrt{n})\|T_0\|}, \tag{3.4}$$

$$\sup_{B_2\times B_1^+} ||I_n - T_0 \cdot D_y F|| \le \frac{\varepsilon}{1+\varepsilon\sqrt{n}}, \quad T_0 = (D_y F(x_0, y_0))^{-1}. \tag{3.5}$$

Then there exists a function $g \in C(B_2; B_1^+)$ satisfying $g(x_0) = y_0$ and

$$F(x, g(x)) \le 0 \quad \forall x \in B_2. \tag{3.6}$$

Moreover the following inequality holds true:

$$|g(x_1) - g(x_2)| \le (1+\varepsilon)||T_0|| \sup_{B_2\times B_1^+} ||D_x F|| \, |x_1 - x_2| \tag{3.7}$$

for every $x_1, x_2 \in B_2$.

Remark 3.5. The inequalities (3.4) and (3.5) are used to derive the size r_2 for the domain of definition and the size r_1 for the set of values of the implicit function. In particular, the first inequality is used to find the radius r_2 as a function of r_1 (i.e., $r_2 = r_2(r_1)$), then from (3.5) one derives the radius r_1.

4. Comments and applications

In [6], the author relates the local upper-Lipschitz property of the solution map (1.2) to the convergence of Newton sequences associated to a system of equations or inequalities.

In this section, we discuss an application of Theorem 3.1 to this topic showing the role played by the estimates of the neighborhoods where the implicit function is defined. Moreover, for the reader's comprehension, we present some examples and applications of Theorem 3.4.

Definition 4.1. A set-valued map Γ from $\mathbb{R}^m$ to the subsets of $\mathbb{R}^n$ is locally upper-Lipschitz continuous at (x^*, y^*), $y^* \in \Gamma(x^*)$, with constants a and b for neighborhoods and c for growth if

$$\Gamma(x) \cap \operatorname{cl} B_a(y^*) \subset y^* + c|x - x^*| \operatorname{cl} B_1(0) \quad \forall x \in \operatorname{cl} B_b(x^*). \tag{4.1}$$

Proposition 4.2. *Let* $\mathbf{S}$ *be the following solution map:*

$$x \in \mathbb{R}^n \longmapsto \mathbf{S}(x) = \{y \in \mathbb{R}^n x = f(y)\}, \tag{4.2}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ *is Lipschitz continuous. Let* $y^* \in \mathbf{S}(x^*)$ *and assume that* $\|(\partial f(y^*))^{-1}\|_\infty < c$*. Let* $\delta_2 = \delta_1/2\nu$*, where* $\nu = \max(1, 2c)$ *and*

$$\delta_1 = \sup\{\rho \geq 0 : \operatorname{co}\{\partial f(B_\rho(y^*))\} \subset \mathcal{S}(\lambda, n)\}. \tag{4.3}$$

Then for any $0 < a < \delta_1$, $0 < b < \delta_2$, $\mathbf{S}$ *is locally upper-Lipschitz continuous at* (x^*, y^*)*, with constants* a *and* b *for neighborhoods and* c *for growth.*

Remark 4.3. With regard to the role of the parameter a in the convergence of Newton sequences, we observe that in [6, Theorem 3.1 and Corollary 3.1], sufficient conditions are given so that any sequence $\{y_n\}$ obtained from the algorithm

$$x = f(y_n) + H_n \cdot (y_{n+1} - y_n), \quad n = 0, 1, 2, \ldots, \tag{4.4}$$

where H_n is a suitable sequence of matrices, and whose elements are all in $B_\sigma(y^*)$, is convergent to y^*. In particular, one of the assumptions for this convergence requires that σ is estimated from above by the constant a. Therefore it is important to have a nontrivial lower bound on a. The amount (4.3) in Proposition 4.2 can be used to this purpose.

Proof of Proposition 4.2. Let $0 < a < \delta_1$ and let $0 < b < \delta_2$. Let $F(x, y) = f(y) - x$, then by the hypotheses, F is Lipschitz continuous and $F(x^*, y^*) = 0$; moreover the assumptions (ii) and (iii) in Theorem 3.1 are satisfied with $\lambda = c$ and $L = 1$. Therefore we can apply this result to F at the reference point (x^*, y^*). By the definition of F, (3.1), and (4.3), we deduce that $r_1 = \delta_1$ and $r_2 = \delta_2$. Hence there exists a unique continuous function $g : B_{\delta_2}(x^*) \to B_{\delta_1}(y^*)$ such that $g(x^*) = y^*$ and $f(g(x)) = x$ for every $x \in B_{\delta_2}(x^*)$.

Let $x \in \operatorname{cl} B_b(x^*)$ and consider $y \in \mathbf{S}(x) \cap \operatorname{cl} B_a(y^*)$, then we have

$$F(x, y) = f(y) - x = 0. \tag{4.5}$$

Remark 3.3 implies that $y = g(x)$, and the inequality (3.3) yields

$$|y - y^*| = |g(x) - g(x^*)| \leq c|x - x^*|. \tag{4.6}$$

Therefore, $y \in y^* + c|x - x^*| \operatorname{cl} B_1(0)$. In light of Definition 4.1, this proves the assertion. □

Remark 4.4. We observe that under the same assumptions, the argument used in the proof of Proposition 4.2 still works in the case of a system of inequalities (i.e., $f(y) - x \leq 0$). In this situation it suffices to consider the map $F(x, y) = [f(y) - x] - [f(y^*) - x^*]$.

In the following example, we construct a lower bound for the supremum in (3.1).

Example 4.5. Let α,β,γ be real-valued, Lipschitz continuous functions on $\mathbb{R}^m$ such that

$$\gamma(0) = 0, \qquad 0 < \alpha(x) < \beta(x) \leq \tau \quad \forall x \in \mathbb{R}^m \tag{4.7}$$

for some positive constant $\tau > 0$. Let

$$F(x,y) = \gamma(x) + \begin{cases} \alpha(x)y, & y \geq 0, \\ \beta(x)y, & y < 0. \end{cases} \tag{4.8}$$

Let $R > 0$, then F is Lipschitz continuous on $\mathbb{R}^n \times (-R,R)$ with a constant estimated by $\sqrt{2}\max(\mathrm{Lip}(\gamma) + MR, \tau)$, where $M = \max(\mathrm{Lip}(\alpha), \mathrm{Lip}(\beta))$. Since $F(0,0) = 0$ and $\partial_y F(0,0) = [\alpha(0),\beta(0)]$, taking $\lambda > 1/\alpha(0)$, we can apply Theorem 3.1 to F at $(0,0)$. We find a lower bound for the radius r_1 defined in (3.1). We choose R and λ to satisfy also the inequality

$$\alpha(0) - R\,\mathrm{Lip}(\alpha) \leq \lambda^{-1}. \tag{4.9}$$

Let $\rho > 0$ be such that

$$\rho\,\mathrm{Lip}(\alpha) < \alpha(0) - \lambda^{-1}. \tag{4.10}$$

It is easy to show that

$$\partial_y F(B_\rho(0) \times (-\rho,\rho)) \subset \bigcup_{|x|<\rho} [\alpha(x),\beta(x)], \tag{4.11}$$

and for any $t \in [\alpha(x),\beta(x)]$, with $|x| < \rho$, it holds that

$$t \geq \alpha(x) \geq \alpha(0) - \rho\,\mathrm{Lip}(\alpha) > \lambda^{-1}. \tag{4.12}$$

Using (4.11) and (4.12), we obtain the inclusion

$$\mathrm{co}\{\partial_y F(B_\rho(0) \times (-\rho,\rho))\} \subset S(\lambda,1). \tag{4.13}$$

Hence, by (3.1), the number ρ belongs to the set whose supremum is r_1, so $r_1 \geq \rho$. By the arbitrary choice of $\rho > 0$ satisfying (4.10), we get the following estimate:

$$r_1 \geq \frac{\alpha(0) - \lambda^{-1}}{\mathrm{Lip}(\alpha)}. \tag{4.14}$$

Remark 4.6. The typical cases where we can apply Theorem 3.4, are the following.

(1) Let $G \in C^1(\mathbb{R}^m \times \mathrm{cl}\,\Omega; \mathbb{R}^n)$, where $\Omega \subset \mathbb{R}^n$ is an open convex set, and consider $x_0 \in \mathbb{R}^m$, $y_0 \in \partial\Omega$, such that $G(x_0, y_0) \le 0$. Then it is easy to show that there exists at least a basis $\mathbf{v} = \{v_1, v_2, \dots, v_n\}$ in $\mathbb{R}^n$ and a positive number r such that the "cones"

$$C_r(y_0;\mathbf{v}) = \left\{ y_0 + \sum_{i=1}^{n} y_i v_i : y_i \in [0,r]\ \forall i = 1,\dots,n \right\},$$
$$C_r^+(y_0;\mathbf{v}) = \left\{ y_0 + \sum_{i=1}^{n} y_i v_i : y_i \in (0,r]\ \forall i = 1,\dots,n \right\} \tag{4.15}$$

satisfy $C_r(y_0;\mathbf{v}) \subset \mathrm{cl}\,\Omega$ and $C_r^+(y_0;\mathbf{v}) \subset \Omega$. For instance, if $\Omega = \{y \in \mathbb{R}^n : y_n < 0\}$ and $y_0 = 0$, then we can consider the system $\mathbf{v} = \{e_1, e_2, \dots, -e_n\}$, $\{e_i\}_{1 \le i \le n}$ being the standard orthonormal basis of $\mathbb{R}^n$. With this choice, the previous inclusions hold for any $r > 0$.

For a basis $\mathbf{v} = \{v_1, v_2, \dots, v_n\}$ satisfying the previous assumptions, consider the function F defined by

$$F(x,y) = G(x, y_0 + V \cdot y), \tag{4.16}$$

where $V \in \mathbb{R}^{n \times n}$ is the matrix whose columns are the vectors $v_1, v_2, \dots, v_n$. Then $F \in C^1(\mathbb{R}^m \times [0,r]^n)$; if $D_y G(x_0, y_0)$ is invertible and the following holds:

$$p \in \mathbb{R}^n, \quad V^{-1} \cdot p \ge 0 \Longrightarrow D_y G(x_0, y_0) \cdot p \le 0, \tag{4.17}$$

then F satisfies the hypotheses of Theorem 3.4 at $(x_0, 0)$: in fact, given the regularity assumptions on G, it is always possible to solve the inequalities (3.4), (3.5) to determine $r_1 \le r$ and r_2. Therefore, there exists a Lipschitz continuous function $g : B_2 \to [0, r_1]^n$ such that $g(x_0) = 0$ and (3.6) holds true. In light of the requirements on $\mathbf{v}$, the Lipschitz continuous function $f = y_0 + V \cdot g$ is $\mathrm{cl}\,\Omega$-valued, and

$$f(x_0) = y_0, \quad G(x, f(x)) \le 0 \quad \forall x \in B_2. \tag{4.18}$$

(2) Let $F : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz continuous function in a neighborhood of a point $(x_0, y_0) \in \mathbb{R}^m \times \mathbb{R}^n$, where $F(x_0, y_0) = 0$. If $\partial_y F(x_0, y_0)$ is not invertible, we cannot apply the implicit function theorem for Lipschitz maps [5]. However, if there exists an open convex set $\Omega \subset \mathbb{R}^n$ such that $y_0 \in \partial\Omega$ and the argument of the previous point can be applied, then we can find an implicit function $g(x)$, defined in a neighborhood of x_0, for the set of inequalities $F \le 0$. We illustrate this situation in the following example.

Example 4.7. Let $F(x,y) = h(x,y) - [y]_+$ where $(x,y) \in \mathbb{R}^2$, $[\cdot]_+$ denotes the positive part of a real number, and h is a smooth function such that $h(0,0) = 0$, $D_y h(0,0) \in [0,1)$. The function F is Lipschitz continuous in a neighborhood of $(0,0)$ where $F(0,0) = 0$, and the generalized partial Jacobian at this point is

$$\partial_y F(0,0) = D_y h(0,0) - [0,1], \tag{4.19}$$

which contains 0. Therefore we cannot apply the Lipschitzian version of the implicit function theorem at $(0,0)$. Nevertheless, by considering the restriction of F to the domain

$\mathbb{R}\times[0,\infty)$, we get $F\in C^1(\mathbb{R}\times[0,\infty))$; moreover, using the convention described in the point (2) of section 2, we have

$$D_yF(0,0)=D_yh(0,0)-1<0. \tag{4.20}$$

Therefore $-D_yF(0,0)$ is invertible and sign preserving, and we can apply Theorem 3.4 to F at $(0,0)$ obtaining an implicit function for the inequality $F\leq 0$.

5. Proof of the results

This section is devoted to the proof of the results presented in Section 3. We proceed by proving Theorem 3.1. To this end we need to state a result which represents a specialized version of [10, Theorem 5.2].

Proposition 5.1. *Let $f:B_r(\xi_0)\subset\mathbb{R}^n\to\mathbb{R}^n$ be a Lipschitz continuous function such that $\operatorname{co}\{\partial f(B_r(\xi_0))\}$ is invertible and there exists $\ell>0$ such that*

$$\|(\operatorname{co}\{\partial f(B_r(\xi_0))\})^{-1}\|_\infty\leq\ell. \tag{5.1}$$

Then the following hold:

$$|f(\xi_0+h)-f(\xi_0)|\geq\frac{1}{\ell}|h|\quad\forall|h|<r, \tag{5.2}$$

$$B_{r/2\ell}(f(\xi_0))\subset f(B_r(\xi_0)). \tag{5.3}$$

Proof of Theorem 3.1. Let $f(x,y)=(F(x,y),x)$ for $(x,y)\in B_r(\xi_0)$ where $\xi_0=(x_0,y_0)\in\mathbb{R}^{m+n}$. We have

$$\partial f(x,y)=\begin{pmatrix}\partial_xF(x,y) & \partial_yF(x,y)\\ I_m & O\end{pmatrix}\quad\forall(x,y)\in B_r(\xi_0), \tag{5.4}$$

where O denotes a null matrix of order $m\times n$. Let $0<\rho<r_1$, then by (3.1) we have the inclusion

$$\operatorname{co}\{\partial_yF(B_\rho(x_0)\times B_\rho(y_0))\}\subset S(\lambda,n). \tag{5.5}$$

By (5.4), if $H\in\operatorname{co}\{\partial f(B_\rho(\xi_0))\}$, then there exist $Q\in\operatorname{co}\{\partial_xF(B_\rho(x_0)\times B_\rho(y_0))\}$, $P\in\operatorname{co}\{\partial_yF(B_\rho(x_0)\times B_\rho(y_0))\}$ such that

$$H=\begin{pmatrix}Q & P\\ I_m & O\end{pmatrix}, \tag{5.6}$$

and by the invertibility of P we have

$$H^{-1}=\begin{pmatrix}O & I_m\\ P^{-1} & -P^{-1}\cdot Q\end{pmatrix}. \tag{5.7}$$

The assumption (iii) yields $\|Q\|\leq L$. Hence, using (5.7) and the definition of $S(\lambda,n)$ in (2.4), we easily get

$$\|H^{-1}\|\leq\max(1,(1+L)\lambda)=\ell. \tag{5.8}$$

We can apply Proposition 5.1 to f obtaining the inclusion

$$B_{\rho/2\ell}(0,x_0) \subset f(B_\rho(x_0,y_0)). \tag{5.9}$$

Let $x \in B_{\rho/2\ell}(x_0)$, then by (5.9) we find $(x',y) \in B_\rho(x_0,y_0)$ such that $f(x',y) = (0,x)$. By the definition of f, $x' = x$ and we set $g(x) = y \in B_\rho(y_0)$. We prove that g is well defined. If there exist y_1, $y_2 \in B_\rho(y_0)$ such that $f(x,y_1) = f(x,y_2) = (0,x)$, then by Theorem 2.4 we have

$$0 = F(x,y_1) - F(x,y_1) \in \text{co}\{\partial_y F(x,[y_1,y_2])\} \cdot (y_1 - y_2). \tag{5.10}$$

Since $\rho/2\ell < \rho < r_1$ and using (5.5), we obtain $y_1 = y_2$. By construction and (i), we get $g(x_0) = y_0$ and the equation (3.2) in $B_{\rho/2\ell}(x_0)$. We show that g is Lipschitz continuous. Let $x_1,x_2 \in B_{\rho/2\ell}(x_0)$ and $g_1 = g(x_1)$, $g_2 = g(x_2)$, then by (3.2) and the mean value theorem, we have

$$\begin{aligned} 0 = F(x_1,g_1) - F(x_2,g_2) &= [F(x_1,g_1) - F(x_1,g_2)] + [F(x_1,g_2) - F(x_2,g_2)] \\ &= P \cdot (g_1 - g_2) + Q \cdot (x_1 - x_2) \end{aligned} \tag{5.11}$$

for some $P \in \text{co}\{\partial_y F(x_1,[g_1,g_2])\}$ and $Q \in \text{co}\{\partial_x F([x_1,x_2],g_2)\}$. Since $[x_1,x_2] \subset B_\rho(x_0)$ and $[g_1,g_2] \subset B_\rho(y_0)$, by (5.5) and (iii), we can write

$$|P \cdot x| \geq \frac{1}{\lambda}|x| \quad \forall x \in \mathbb{R}^n, \qquad \|Q\| \leq L. \tag{5.12}$$

Using (5.12) in (5.11), we get

$$|g_1 - g_2| \leq \lambda |P \cdot (g_1 - g_2)| = \lambda |Q \cdot (x_1 - x_2)| \leq \lambda L |x_1 - x_2|. \tag{5.13}$$

For any fixed $\rho \in (0,r_1)$, the relation (5.10) implies the uniqueness of g. The previous analysis proves that the following set is nonempty:

$$\begin{aligned} X = \Big\{(g,\rho) : \rho \leq r_1,\ g \in C(B_{\rho/2\ell}(x_0);\ B_\rho(y_0)),\ g(x_0) = y_0,\ g \text{ satisfies (3.2)} \\ \text{and (3.3) in } B_{\rho/2\ell}(x_0)\Big\}. \end{aligned} \tag{5.14}$$

We assign the following order relation $(\leq_X)$ on X:

$$(g,\rho) \leq_X (h,\delta) \Longleftrightarrow \rho \leq \delta, \tag{5.15}$$

and h extends g. By a standard argument based on the Zorn lemma, we deduce that X admits a maximal element with respect to the relation $\leq_X$. Let $(\hat{g},\hat{\rho}) \in X$ be such element, then $\hat{\rho} = r_1$. Otherwise, for a fixed $\hat{\rho} < \rho_1 < r_1$, we can repeat the previous construction to find a function $g_1 \in C(B_{\rho_1/2\ell}(x_0);\ B_{\rho_1}(y_0))$ satisfying $g(x_0) = y_0$, (3.2), and (3.3) over $B_{\rho_1/2\ell}(x_0)$. Since $(g_1,\rho_1) \in X$ and g_1 is unique on its domain of definition, we have $(\hat{g},\hat{\rho}) \leq_X (g_1,\rho_1)$. Since $\hat{\rho} < \rho_1$, this yields a contradiction proving that $\hat{\rho} = r_1$. Hence the maximal element of X is the implicit function we are seeking. This concludes the proof. □

Proof of Proposition 5.1. Applying the mean value Theorem 2.4, we have

$$f(\xi_0+h) - f(\xi_0) \in \operatorname{co}\{\partial f(\xi_0+th) : t \in [0,1]\} \cdot h \quad \forall |h| < r, \tag{5.16}$$

and by the hypotheses, for every $S \in \operatorname{co}\{\partial f(B_r(\xi_0))\}$, we have $|Sx| \geq \ell^{-1}|x|$ for any $x \in \mathbb{R}^n$, therefore (5.2) follows. To prove the inclusion (5.3), consider $r > \rho > 0$ and $\eta \in B_{\rho/2\ell}(\eta_0)$, η_0 being $f(\xi_0)$. Let $\Phi(\xi) = |\eta - f(\xi)|^2$. This function takes its minimum over $\operatorname{cl} B_\rho(\xi_0)$ in a point $\bar{\xi}$. In fact $\bar{\xi} \in B_\rho(\xi_0)$, otherwise we have $|\bar{\xi} - \xi_0| = \rho$ and by (5.2) we get

$$|\eta - f(\xi_0)| \geq |f(\xi_0) - f(\bar{\xi})| - |\eta - f(\bar{\xi})| \geq \frac{\rho}{\ell} - |\eta - f(\xi_0)| \tag{5.17}$$

implying $|\eta - \eta_0| \geq \rho/2\ell$, which is false of course. Therefore we conclude that $\bar{\xi}$ lies in the interior of the ball. Now if $\eta = f(\bar{\xi})$, we are done. Otherwise, by the optimality condition [9], it holds that

$$0 \in 2\big(f(\bar{\xi}) - \eta\big)^{\top} \cdot \partial f(\bar{\xi}), \tag{5.18}$$

where $(\cdot)^{\top}$ denotes the transpose. Since $\partial f(\bar{\xi})$ contains only invertible matrices, (5.18) is true only when $f(\bar{\xi}) - \eta = 0$, and this is a contradiction. Therefore we have proved the following inclusion:

$$B_{\rho/2\ell}(\eta_0) \subset f(B_\rho(\xi_0)) \tag{5.19}$$

for any $0 < \rho < r$. By the arbitrary choice of ρ, we obtain the inclusion (5.3). □

Proof of Theorem 3.4. Let $G(\cdot,\cdot) = F(\cdot,\cdot) - F(x_0,y_0)$, then we use a fixed-point argument applied to the following map:

$$\begin{gathered}\Phi : C(B_2;B_1^+) \longrightarrow C(B_2;B_1^+), \\ \Phi(v)(x) = y_0 + \big[v(x) - T_0 \cdot G(x,v(x)) - y_0\big]_+ \quad \forall v \in C(B_2;B_1^+),\ x \in B_2,\end{gathered} \tag{5.20}$$

where $T_0 = (D_yF(x_0,y_0))^{-1}$ and $[\xi]_+$ denotes the vector whose entries are the positive parts of the components of ξ. We consider the supremum norm on $C(B_2;B_1^+)$. Using the assumption $(\jmath\jmath\jmath\jmath)$, we show that Φ is well defined and is a contraction. Let $v, v' \in C(B_2;B_1^+)$, $x \in B_2$; since the Lipschitz constant of $\tau \in \mathbb{R} \mapsto [\tau]_+$ is 1 and B_1^+ is convex, by $(\jmath\jmath)$ and (3.5) we have

$$\begin{aligned}|\Phi(v)(x) - \Phi(v')(x)| &\leq |v(x) - T_0 \cdot G(x,v(x)) - v'(x) + T_0 \cdot G(x,v'(x))| \\ &\leq \sup_{B_2 \times B_1^+} \|I_n - T_0 \cdot D_yF\| \, |v(x) - v'(x)| \\ &\leq \frac{\varepsilon}{1+\varepsilon\sqrt{n}} |v(x) - v'(x)|.\end{aligned} \tag{5.21}$$

Therefore Φ is a contraction map. Let $\overline{v}$ be the constant function y_0, then by $\Phi(\overline{v})(x_0) = y_0$, (3.4), and (5.21), we can write

$$\begin{aligned} 0 &\leq (\Phi(v)(x) - y_0)_i \leq |\Phi(v)(x) - y_0| \\ &\leq |\Phi(v)(x) - \Phi(\overline{v})(x)| + |\Phi(\overline{v})(x) - \Phi(\overline{v})(x_0)| \\ &\leq \frac{\varepsilon\sqrt{n}}{(1+\varepsilon\sqrt{n})} r_1 + ||T_0|| \sup_{x\in B_2} |F(x,y_0) - F(x_0,y_0)| \\ &\leq \frac{\varepsilon\sqrt{n}}{1+\varepsilon\sqrt{n}} r_1 + ||T_0|| \frac{1}{1+\varepsilon\sqrt{n}} \frac{r_1}{||T_0||} = r_1 \quad \forall i = 1,\dots,n. \end{aligned} \tag{5.22}$$

This yields $\Phi(v)(x) \in B_1^+$ and it proves that Φ is well defined. Since $C(B_2;B_1^+)$ is a closed subspace of the Banach space $C(B_2;\mathbb{R}^n)$, we can apply the fixed-point theorem getting a unique function $g \in C(B_2;B_1^+)$ such that

$$g(x) = y_0 + [g(x) - T_0 \cdot G(x,g(x)) - y_0]_+ \quad \forall x \in B_2. \tag{5.23}$$

Let $x \in B_2$ and $1 \leq i \leq n$; if $(g(x) - y_0)_i = 0$, then (5.14) implies

$$(T_0 \cdot G(x,g(x)))_i \geq 0. \tag{5.24}$$

Otherwise, if $(g(x) - y_0)_i > 0$, then $(g(x) - T_0 \cdot G(x,g(x)) - y_0)_i > 0$, and by (5.23) we have

$$(T_0 \cdot G(x,g(x)))_i = 0. \tag{5.25}$$

In conclusion $p = T_0 \cdot G(x,g(x)) \geq 0$. Using $(\jmath\jmath\jmath)$, the matrix $-T_0^{-1}$ is sign preserving, hence

$$G(x,g(x)) = T_0^{-1} \cdot p \leq 0 \quad \forall x \in B_2. \tag{5.26}$$

By $(\jmath)$ and the definition of G, we obtain (3.6).

In the fixed-point theorem, the function g is obtained as the limit of the following sequence:

$$g_{n+1} = \Phi(g_n) \quad \forall n \geq 0. \tag{5.27}$$

Choosing $g_0 = \overline{v}$ and using $G(x_0,y_0) = 0$, it is easy to show, by induction, that $g_n(x_0) = y_0$ for every n, which yields $g(x_0) = y_0$. Finally, we prove the Lipschitz continuity of g. Let $x_1,x_2 \in B_2$ and $g_1 = g(x_1)$, $g_2 = g(x_2)$, then by (5.23) and (3.5) we have

$$\begin{aligned} |g_1 - g_2| &\leq |g_1 - T_0 \cdot G(x_1,g_1) - g_2 + T_0 \cdot G(x_2,g_2)| \\ &\leq |g_1 - T_0 \cdot G(x_1,g_1) - g_2 + T_0 \cdot G(x_1,g_2)| + ||T_0|| |G(x_1,g_2) - G(x_2,g_2)| \\ &\leq \frac{\varepsilon}{1+\varepsilon} |g_1 - g_2| + ||T_0|| \sup_{B_2\times B_1^+} ||D_xF|| |x_1 - x_2|. \end{aligned} \tag{5.28}$$

The previous chain of inequalities proves (3.7). □

References

[1] R. B. Ash, *Complex Variables*, Academic Press, New York, 1971.

[2] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed., Athena Scientific, Massachusetts, 1999.

[3] H. C. Chang, W. He, and N. Prabhu, *The analytic domain in the implicit function theorem*, JIPAM. J. Inequal. Pure Appl. Math. **4** (2003), no. 1, Article 12, 1–5.

[4] F. H. Clarke, *On the inverse function theorem*, Pacific J. Math. **64** (1976), no. 1, 97–102.

[5] ———, *Optimization and Nonsmooth Analysis*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley & Sons, New York, 1983.

[6] A. L. Dontchev, *Lipschitzian stability of Newton's method for variational inclusions*, System Modelling and Optimization (Cambridge, 1999), Kluwer Academic Publishers, Massachusetts, 2000, pp. 119–147.

[7] R. Fletcher, *Practical Methods of Optimization*, 2nd ed., Wiley-Interscience, New York, 2001.

[8] L. Hörmander, *An Introduction to Complex Analysis in Several Variables*, 2nd revised ed., North-Holland Mathematical Library, vol. 7, North-Holland Publishing, Amsterdam, 1973.

[9] V. Jeyakumar and D. T. Luc, *Nonsmooth calculus, minimality, and monotonicity of convexificators*, J. Optim. Theory Appl. **101** (1999), no. 3, 599–621.

[10] ———, *An open mapping theorem using unbounded generalized Jacobians*, Nonlinear Anal. Ser. A: Theory Methods **50** (2002), no. 5, 647–663.

[11] S. G. Krantz and H. R. Parks, *The Implicit Function Theorem. History, Theory, and Applications*, Birkhäuser Boston, Massachusetts, 2002.

[12] Z.-Q. Luo, J.-S. Pang, and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[13] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optimization **15** (1977), no. 6, 959–972.

[14] B. Mordukhovich, *Lipschitzian stability of constraint systems and generalized equations*, Nonlinear Anal. **22** (1994), no. 2, 173–206.

[15] ———, *Sensitivity analysis for constraint and variational systems by means of set-valued differentiation*, Optimization **31** (1994), no. 1, 13–46.

[16] S. Nash and A. Sofer, *Linear and Nonlinear Programming*, McGraw-Hill, New York, 1995.

[17] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Series in Operations Research, Springer-Verlag, New York, 1999.

[18] L. Q. Qi and J. Sun, *A nonsmooth version of Newton's method*, Math. Programming **58** (1993), no. 3, Ser. A, 353–367.

[19] S. M. Robinson, *Generalized equations and their solutions. I. Basic theory*, Math. Programming Stud. (1979), no. 10, 128–141.

Marco Papi: Istituto per le Applicazioni del Calcolo "Mauro Picone," viale del Policlinico 137, 00161 Roma, Italy; Dipartimento di Matematica, Università degli Studi di Roma "Tor Vergata," via della Ricerca Scientifica, 00133 Roma, Italy

E-mail addresses: papi@iac.rm.cnr.it; ivnpp@tiscalinet.it