**Control and Cybernetics**

vol. **32** (2003) No. 3

# Robinson's implicit function theorem

by

**Asen L. Dontchev**

Mathematical Reviews
Ann Arbor, MI 48107-8604
ald@ams.org

**Abstract:** Robinson's implicit function theorem has played a mayor role in the analysis of stability of optimization problems in the last two decades. In this paper we take a new look at this theorem, and with an updated terminology go back to the roots and present some extensions.

**Keywords:** implicit function, sensitivity, variational inequality, optimization.

## 1. Robinson's theorem

Given two sets $X$ and $Y$, we denote by $F : X \rightrightarrows Y$ a *set-valued* mapping $F$ acting from $X$ to the subsets of $Y$. The *domain* of $F$ is defined as $\operatorname{dom} F = \{x \in X \mid F(x) \neq \emptyset\}$ while its *range* as $\operatorname{rge} F = \{y \mid y \in F(x), x \in \operatorname{dom} F\}$. The *graph* of a mapping $F : X \rightrightarrows Y$ is $\operatorname{gph} F = \{(x, y) \in X \times Y \mid y \in F(x)\}$. If for any $x \in \operatorname{dom} F$, $F(x)$ consists of exactly one element, we say that $F$ is a function or an operator and write $F : X \to Y$. The inverse of a mapping $F : X \rightrightarrows Y$ is another mapping, denoted $F^{-1}$, and defined as

$$Y \ni y \mapsto F^{-1}(y) = \{x \in X \mid y \in F(x)\};$$

that is,

$$y \in F(x) \quad \Leftrightarrow \quad x \in F^{-1}(y).$$

Definition 1.1 (graphical localization) *Given a mapping $F$ acting from a topological space $X$ to the subsets of a topological space $Y$ with $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, a graphical localization of $F$ around $(\bar{x}, \bar{y})$ is a mapping $\tilde{F}$ whose graph is the graph of $F$ restricted to a "box" $U \times V$, for some neighborhoods $U$ of $\bar{x}$ and $V$ of $\bar{y}$; that is,*

$$\operatorname{gph} \tilde{F} = \operatorname{gph} F \cap (U \times V).$$

In other words, a graphical localization of a mapping $F$ around $(\bar{x}, \bar{y})$ is

$$\tilde{F}(x) = \begin{cases} F(x) \cap V & \text{when } x \in U, \\ \emptyset & \text{otherwise.} \end{cases}$$

The inverse of $\tilde{F}$ satisfies

$$\tilde{F}^{-1}(y) = \begin{cases} F^{-1}(y) \cap U & \text{when } y \in V, \\ \emptyset & \text{otherwise,} \end{cases}$$

and hence $\tilde{F}^{-1}$ is a graphical localization of $F^{-1}$ around $(\bar{y}, \bar{x})$. The domain of a graphical localization $\tilde{F}$ may be different from $\operatorname{dom} F \cap V$ and depends on the choice of $U$.

The classical implicit function theorem is about mappings that are implicitly defined by equations; that is, mappings of the form:

$$P \ni p \mapsto \{x \in X \mid f(p, x) = 0\}, \tag{1}$$

where $P$ is the space of "parameters" $p$ and $f : P \times X \to Y$. Assuming that $P$, $X$ and $Y$ are, e.g., Banach spaces and $f : P \times X \to Y$ with $(\bar{p}, \bar{x}) \in \operatorname{int} \operatorname{dom} f$ is a continuously Fréchet differentiable ($\mathcal{C}^1$) function in a neighborhood of $(\bar{p}, \bar{x})$, it claims that if the partial derivative of $f$ with respect to $x$ at $(\bar{p}, \bar{x})$, $\nabla_x f(\bar{p}, \bar{x})$, is an invertible operator, then the mapping (1) has a single-valued graphical localization $x(\cdot)$ around $(\bar{p}, \bar{x})$ which is $\mathcal{C}^1$. Furthermore, the derivative of the implicit function $p \mapsto x(p)$ can be computed by using the chain rule.

In a landmark paper, Robinson (1980), S. M. Robinson proved an implicit function theorem for variational inequalities that goes beyond the format of the classical theory. Let $X$ be a Banach space and $X^*$ be its dual. For a mapping $f : X \to X^*$ and a nonempty convex closed set $C \subset X$, the variational inequality problem is as follows:

$$\text{Find } x \in C \text{ such that } \quad \langle f(x), v - x \rangle \geq 0 \quad \text{for all} \quad v \in C. \tag{2}$$

In terms of the normal cone mapping defined as

$$X \ni x \mapsto N_C(x) = \begin{cases} \{\, y \in X^* \mid \langle y, v - x \rangle \leq 0 \text{ for all } v \in C \} & \text{for } x \in C, \\ \emptyset & \text{otherwise,} \end{cases}$$

the variational inequality (2) can be written as the inclusion

$$f(x) + N_C(x) \ni 0 \tag{3}$$

which Robinson called *generalized equation.* For $C = X$, the normal cone $N_X(x) = \{0\}$ for all $x \in X$ and we come to the equation $f(x) = 0$.

To put the stage for Robinson's theorem, we make (2) dependent on a parameter $p$ from a topological space $P$:

$$\langle f(p, x), v - x \rangle \geq 0 \quad \text{for all} \quad v \in C, \tag{4}$$

where now $f : P \times X \to X^*$. Let $\bar{x}$ be a solution of (4) for the value $\bar{p}$ of the parameter, that is, $0 \in f(\bar{p}, \bar{x}) + N_C(\bar{x})$. Next come the two assumptions in Robinson's theorem. The first one concerns the smoothness of $f$:

(R1) *$f$ is Fréchet differentiable with respect to $x$ in a neighborhood of $(\bar{p}, \bar{x})$ and both $f$ and $\nabla_x f$ are continuous in this neighborhood.*

The second condition is about an auxiliary variational inequality involving the *linearization* with respect to $x$ of the function $f$. Robinson called this property of the variational inequality (2) *strong regularity*:

(R2) *The mapping*

$$(f(\bar{p}, \bar{x}) + \nabla_x f(\bar{p}, \bar{x})(\cdot - \bar{x}) + N_C(\cdot))^{-1} \tag{5}$$

*has a single-valued localization around $(0, \bar{x})$ which is Lipschitz continuous near zero.*

The original formulation of Robinson's theorem is as follows:

Theorem 1.1 (Robinson (1980) *Let the variational inequality (5) satisfy the conditions (R1) and (R2) and let $L$ be a Lipschitz constant of the single-valued graphical localization of the mapping (5). Then, for every $\varepsilon > 0$ there exist neighborhoods $U$ of $\bar{x}$ and $V$ of $\bar{p}$ and a function $x : V \to U$ such that for each $p \in V$, $x(p)$ is the unique solution of the variational inequality (4) in $U$ and also for every $p, q \in V$ one has*

$$\|x(p) - x(q)\| \le (L + \varepsilon)\|f(p, x(p)) - f(q, x(p))\|. \tag{6}$$

In other words, under (R1) and (R2) the mapping

$$p \mapsto \{x \in X \mid x \text{ is a solution of (4) for } p\ \} \tag{7}$$

has a lipschitzian single-valued localization around $(\bar{p}, \bar{x})$. Moreover, if (R2) holds with a Lipschitz constant $L$, then for any $\varepsilon$ there exists a single-valued localization of (7) that satisfies (6). An immediate corollary of Robinson's theorem is

Corollary 1.1 *In addition to (R1) and (R2), let $P$ be a metric space with a metric $d_P(\cdot, \cdot)$ and let the function $f$ be Lipschitz continuous with respect to $p$ near $\bar{p}$ uniformly in $x$ near $\bar{x}$; that is, there exists a constant $l$ such that for every $p, q$ near $\bar{p}$ and for every $x$ near $\bar{x}$*

$$\|f(p, x) - f(q, x)\| \le l d_P(p, q).$$

*Then, for any $\varepsilon > 0$ the mapping (7) has a lipschitzian single-valued graphical localization around $(\bar{p}, \bar{x})$ with a Lipschitz constant $(L + \varepsilon)\, l$.*

While the central role in the classical implicit function theorem is given to the (continuous) differentiability, in Robinson's theorem the main player is the property of Lipschitz continuity, and this is apparently the best one can get for variational problems with constraints, in general. The good news is that with Lipschitz continuity one can get quite a lot, and in particular prove convergence of optimization algorithms and estimate errors of approximations of variational problems.

Numerous applications of Robinson's theorem in optimization are available in the literature, see e.g. the recent books by Bonnans and Shapiro (2000) and Klatte and Kummer (2002). Robinson's theorem has also found many applications to optimal control problems for both ordinary and partial differential equations, for an extended state-of-the-art review see Malanowski (2001a).

## 2. Extensions

We adopt some terminology from Rockafellar and Wets (1997). Let $X$ and $Y$ be metric spaces with metrics $d_X(\cdot)$ and $d_Y(\cdot)$, respectively. Recall that the *Lipschitz modulus* of a function $f : X \to Y$ at a point $\bar{x} \in \operatorname{int} \operatorname{dom} f$ is

$$\operatorname{lip} f(\bar{x}) := \limsup_{\substack{x,x' \to \bar{x} \\ x \neq x'}} \frac{d_Y(f(x'), f(x))}{d_X(x', x)}.$$

If $f$ is Lipschitz continuous in a set $D \subset X$ with a constant $L$ then lip $f(\bar{x}) \le L$ for every $\bar{x} \in D$. Conversely, if there exist a constant $L$ and a neighborhood of $U$ of $\bar{x}$ such that $f$ is Lipschitz continuous in $U$ with a constant $L$, then $\operatorname{lip} f(\bar{x}) \le L$; the absence of this property is identified by $\operatorname{lip} f(\bar{x}) = \infty$. For a topological space $P$ and a function $h : P \times X \to Y$, the *partial Lipschitz modulus* with respect to $x$ at $(\bar{p}, \bar{x}) \in \operatorname{dom} h$ is defined as

$$\operatorname{lip}_x h(\bar{p}, \bar{x}) := \lim sup_{\substack{x,x' \to \bar{x}, p \to \bar{p} \\ x \neq x'}} \frac{d_Y(h(p, x'), h(p, x))}{d_X(x', x)}.$$

For a set-valued mapping $S : X \rightrightarrows Y$ with $(\bar{x}, \bar{y}) \in \operatorname{gph} S$, we use the following convention: if $S$ has a single-valued graphical localization around $(\bar{x}, \bar{y})$, with appropriate neighborhoods $U$ of $\bar{x}$ and $V$ of $\bar{y}$, which is Lipschitz continuous in $U$, we denote

$$\operatorname{lip} S(\bar{x} \,|\, \bar{y}) = \operatorname{lip}(S \cap V)(\bar{x}),$$

where $S \cap V : U \to V$ is the associate single-valued graphical localization. The absence of a single-valued graphical localization of $S$ at $(\bar{x}, \bar{y})$ or the case when the single-valued graphical localization of $S$ at $(\bar{x}, \bar{y})$ is not Lipschitz continuous near $\bar{x}$, is signaled by $\operatorname{lip} S(\bar{x} \,|\, \bar{y}) = \infty$. For a function $s$ we have $\operatorname{lip} s(\bar{x}) = \operatorname{lip} s(\bar{x} \,|\, s(\bar{x}))$.

DEFINITION 2.1 (strong regularity) *A mapping $F : X \rightrightarrows Y$ is strongly regular at $\bar{x}$ for $\bar{y}$ if $F^{-1}$ has a lipschitzian single-valued graphical localization around $(\bar{y}, \bar{x})$.*

In other words,

$$F \text{ is strongly regular at } \bar{x} \text{ for } \bar{y} \quad \Leftrightarrow \quad \operatorname{lip} F^{-1}(\bar{y}\,|\,\bar{x}) < \infty.$$

The Lipschitz modulus of the inverse $F^{-1}(\bar{y}\,|\,\bar{x})$ is also called *strong regularity modulus*. If $F$ is a function and strongly regular at $\bar{x}$ for $\bar{y} = F(\bar{x})$, we simply say $F$ is strongly regular at $\bar{x}$.

The simplest example of a strongly regular mapping is any invertible operator $A \in \mathcal{L}(X, Y)$ acting, e.g., in Banach spaces $X$ and $Y$. Indeed, if $A$ is invertible, then $A^{-1}$ is single-valued everywhere and $\operatorname{lip} A^{-1} = \|A^{-1}\|$. From the classical inverse function theorem, a function $f : X \to Y$, which is $\mathcal{C}^1$ near $\bar{x}$, is strongly regular at $\bar{x}$ if and only if $\nabla f(\bar{x})$ is invertible and then $\operatorname{lip} f^{-1}(\bar{y}\,|\,\bar{x}) = \|\nabla f(\bar{x})^{-1}\|$. We use this "double" terminology for convenience only: in some cases it is easier to operate with the mapping itself than with its inverse. E.g., it is more convenient to say "an invertible matrix $A$" than "a matrix $A$ whose inverse mapping $A^{-1}$ is a matrix."

One should note that in his original definition Robinson (1980) called the variational inequality (2) strongly regular when the inverse of the linearization (5) had a lipschitzian single-valued graphical localization. Here we use the term "strong regularity" as a property of a general set-valued mapping. This definition is equivalent to Robinson's definition for the mappings Robinson had in mind.

Robinson's theorem (1.1) will be deduced from the following more general result which, in turn, can be extracted from the original proof of Robinson (1980).

THEOREM 2.1 (Extended Robinson's Theorem) *Let $X$ be a complete metric space, $Y$ a metric space, and $P$ a topological space. Consider a mapping $\mathbf{M} : Y \rightrightarrows X$ with $(\bar{y}, \bar{x}) \in \operatorname{gph} \mathbf{M}$ and a function $h : P \times X \to Y$ which is continuous at $(\bar{p}, \bar{x})$ and with $\bar{y} = h(\bar{p}, \bar{x})$. If $\operatorname{lip} \mathbf{M}(\bar{y}\,|\,\bar{x}) < \gamma$ and $\operatorname{lip}_x h(\bar{p}, \bar{x}) < \lambda < \gamma^{-1}$, then the mapping*

$$P \ni p \mapsto \mathbf{N}(p) := \left\{ x \in X \,\middle|\, x \in \mathbf{M}(h(p, x)) \right\}$$

*has a lipschitzian single-valued graphical localization $x(\cdot)$ around $(\bar{p}, \bar{x})$ which satisfies*

$$d_X(x(p), x(q)) < (\gamma^{-1} - \lambda)^{-1} d_Y(h(p, x(p)), h(q, x(p))) \tag{8}$$

*for all $p, q$ near $\bar{p}$.*

*Proof.* Choose neighborhoods $U$ of $\bar{x}$, $V$ of $\bar{y}$ and $Q$ of $\bar{p}$ such that $m := \mathbf{M} \cap U$ is a function which is Lipschitz continuous on $V$ with a Lipschitz constant $\gamma$ and $h$

is Lipschitz continuous in $x \in U$ with a Lipschitz constant $\lambda$ for all $p \in Q$. Also, take $U$ and $Q$ smaller if necessary so that $h(p,x) \in V$ for all $(p,x) \in Q \times U$. Choose a positive number $a$ such that $I\!B_a(\bar{x})$, the closed ball centered at $\bar{x}$ with radius $a$, is a subset of $U$. Also, chose a number $\delta$ with $0 < \delta < a(1-\gamma\lambda)$ and take $Q$ smaller, if necessary, to obtain that

$$d_Y(h(p,\bar{x}), h(\bar{p},\bar{x})) \le \delta \qquad \text{for all } p \in Q.$$

Fix $p \in Q$ and consider the mapping

$$I\!B_a \ni x \mapsto \Phi_p(x) := m(h(p,x)).$$

Since $\bar{x} = m(h(\bar{p},\bar{x}))$, for any $x \in I\!B_a(\bar{x})$ we have

$$\begin{aligned} d_X(\Phi_p(x), \bar{x}) &= d_X(m(h(p,x)), m(h(\bar{p},\bar{x}))) \\ &\le \gamma(d_Y(h(p,x), h(p,\bar{x})) + d_Y(h(p,\bar{x}), h(\bar{p},\bar{x}))) \\ &\le \gamma(\lambda d_X(x,\bar{x}) + \delta) \le \gamma\lambda a + \delta \le a. \end{aligned}$$

Further, for every $x', x'' \in I\!B_a(\bar{x}), x' \neq x''$,

$$\begin{aligned} & d_X(\Phi_p(x'), \Phi_p(x'')\| = d_X m(h(p,x')), m(h(p,x''))) \\ & \qquad \le \gamma d_Y(h(p,x'), h(p,x'')) \le \gamma\lambda d_X(x',x'') < d_X(x',x''). \end{aligned}$$

Thus, by the contracting mapping theorem, $\Phi_p$ is a unique fixed point $x(p)$ for every $p \in Q$. This means that the mapping $\mathbf{N}$ has a single-valued graphical localization $p \mapsto x(p)$ around $(\bar{p},\bar{x})$. Moreover, since $x(p) = m(h(p,x(p)))$, for any $p,q \in Q$,

$$\begin{aligned} d_X(x(p), x(q)) &\le d_X(m(h(p,x(p))), m(h(q,x(q)))) \\ &\le \gamma(d_Y(h(p,x(p)), h(q,x(p))) + \lambda d_X(x(p),x(q))); \end{aligned}$$

that is, $x$ satisfies (8) and the proof is complete. ■

*Proof of Theorem 1.2* Apply 2.1 with

$$\mathbf{M} = \Big(f(\bar{p},\bar{x}) + \nabla_x f(\bar{p},\bar{x})(\cdot - \bar{x}) + N_C(\cdot)\Big)^{-1}$$

and

$$h(p,x) = f(p,x) - f(\bar{p},\bar{x}) - \nabla_x f(\bar{x},\bar{p})(x - \bar{x}),$$

and with $\bar{y} = 0$. Then $\mathbf{N}$, as defined in 2.1, is exactly the mapping from (7), that is, the solution mapping of the variational inequality (4). The condition (R1) implies that $h$ is continuous at $(\bar{p},\bar{x})$ and $\mathrm{lip}_x\, h(\bar{p},\bar{x}) = 0$ while (R2) means that $\mathbf{M}$ has a lipschitzian single-valued localization around $(0,\bar{x})$. To complete the proof it is sufficient to note that for $\gamma = L$ and any $\varepsilon > 0$ one can choose $\lambda > 0$ so small that $L + \varepsilon \ge (L^{-1} - \lambda)^{-1}$. ■

While the strikingly simple proof of 2.1 above is close to the original proof of Robinson, the result itself goes beyond the framework of the variational inequalities. In particular, Robinson's theorem can be stated for inclusions of the form

$$f(p,x) + F(x) \ni 0,$$

where the normal cone mapping $N_C$ is replaced by *any set-valued mapping* $F : X \rightrightarrows Y$. For the function $f$ it is sufficient to require $f$ have a first-order approximation $g$ such that $g+F$ is strongly regular. Specifically, by a first-order approximation we mean the following.

Definition 2.2 *Let $X$ be a metric space, $Y$ be a linear normed space, and $P$ be a topological space. For $f : P \times X \to Y$ with $(\bar p, \bar x) \in \operatorname{int}\operatorname{dom} f$, a function $g : X \to Y$ with $\bar x \in \operatorname{int}\operatorname{dom} g$ is said to be a first-order approximation of $f$ at $(\bar p, \bar x)$ when, for $\varphi(p,x) = f(p,x) - g(x)$ one has*

$$\operatorname{lip}_x \varphi(\bar p, \bar x) = 0.$$

This concept was introduced by Robinson (1991) where he called it *strong approximation.* For example, for $X$ and $Y$ Banach spaces, if $f$ is Fréchet differentiable in $x$ around $(\bar p, \bar x)$ and its partial derivative $\nabla_x f$ is continuous near $(\bar p, \bar x)$, then $x \mapsto \nabla_x f(\bar p, \bar x)x$ is a first-order approximation of $f$ at $(\bar p, \bar x)$. Moreover, for $h(p,x) = f(p,x) + \varphi(x)$, where $f$ is Fréchet differentiable in $x$ around $(\bar p, \bar x)$ and its partial derivative $\nabla_x f$ is continuous near $(\bar p, \bar x)$, and $\varphi$ is *any* function, the function $x \mapsto \nabla_x f(\bar p, \bar x)x + \varphi(x)$ is a first-order approximation of $f + \varphi$ at $(\bar p, \bar x)$. If $f$ is not dependent on the parameter $p$, then $g$ is a first-order approximation of $f$ (and $f$ is a first-order approximation of $g$) at $\bar x$ if and only if $\operatorname{lip}(f - g)(\bar x) = 0$ which is the same to say that the strict derivative of the difference $f - g$ at $\bar x$ is zero. Putting all this together we obtain

Theorem 2.2 *Let $X$ be a complete metric space, $Y$ a linear normed space, and $P$ a topological space. Given $F : X \rightrightarrows Y$ and a function $f : P \times X \to Y$ with $\bar y \in f(\bar p, \bar x) + F(\bar x)$ and such that $f(\cdot, \bar x)$ is continuous at $\bar p$, let $g : X \to Y$ be a first-order approximation of $f$ at $(\bar p, \bar x)$ and let $g + F$ be strongly regular at $\bar x$ for $\bar y$ with associated Lipschitz constant $L$ of the graphical localization of $(g+F)^{-1}$ around $(\bar y, \bar x)$. Then the mapping*

$$p \mapsto \{x \in X \mid f(p,x) + F(x) \ni 0\} \tag{9}$$

*has a single-valued localization $x(\cdot)$ around $(\bar p, \bar x)$ with the property that for every $\varepsilon > 0$ the inequality (6) holds for $p$ sufficiently close to $\bar p$.*

*Proof.* We apply 2.1 with $\mathbf{M} = (g+F)^{-1}$ and $h(p,x) = -f(p,x) + g(x)$. Then for $(p,x)$ close to $(\bar p, \bar x)$,

$$\begin{aligned} \|h(p,x) - h(\bar p, \bar x)\| &\le \|f(p,x) - g(x) - f(p,\bar x) + g(\bar x)\| \\ &\quad + \|f(p,\bar x) - f(\bar p, \bar x)\| \\ &\le \varepsilon d_X(x, \bar x) + \|f(p, \bar x) - f(\bar p, \bar x) \end{aligned}$$

from which it follows that $h$ is continuous at $(\bar{p}, \bar{x})$. Moreover, from the definition of the strong approximation, $\operatorname{lip}_x h(\bar{p}, \bar{x}) = 0$. The mapping $\mathbf{N}$ is the one in (9) and hence 2.1 applies with $\bar{y} = 0$. ■

Observe that 2.1 is not really about approximation or linearization of a function. It is about a relation between two constants, the modulus of the lipschitzian single-valued localization of a given mapping and the Lipschitz modulus of a perturbing function. This observation is becoming more explicit in the following result which is obtained from 2.1. In the next two assertions the spaces $X$ and $Y$ are as in 2.2.

THEOREM 2.3 *Consider a set-valued mapping* $\mathbf{F} : X \rightrightarrows Y$ *and a point* $(\bar{x}, \bar{y}) \in \operatorname{gph} \mathbf{F}$. *Consider also a function* $G : X \to Y$ *with* $\bar{x} \in \operatorname{dom} G$. *If* $\operatorname{lip} \mathbf{F}^{-1}(\bar{y}\,|\,\bar{x}) \cdot \operatorname{lip} G(\bar{x}) < 1$, *then*

$$\operatorname{lip}(\mathbf{F} + G)^{-1}(\bar{y} + G(\bar{x})\,|\,\bar{x}) \quad \leq \quad (\operatorname{lip} \mathbf{F}^{-1}(\bar{y}\,|\,\bar{x})^{-1} - \operatorname{lip} G(\bar{x}))^{-1}.$$

*Proof.* Apply 2.1 with $h(p, x) = p - G(x)$ and $\mathbf{M} = \mathbf{F}^{-1}$. We have

$$\mathbf{N}(p) = \{x \in X \mid x \in \mathbf{F}^{-1}(p - G(x))\} = (\mathbf{F} + G)^{-1}(p)$$

and $\operatorname{lip}_x h(\bar{p}, \bar{x}) = \operatorname{lip} G(\bar{x})$. Also, $h(p, x(p)) - h(q, x(p)) = p - q$. It remains to substitute $\bar{y} = \bar{p}$. ■

In particular, we have

COROLLARY 2.1 *Consider a set-valued mapping* $\mathbf{F} : X \rightrightarrows Y$ *with* $\operatorname{lip} \mathbf{F}^{-1}(\bar{y}\,|\,\bar{x}) < \infty$ *and a function* $G : X \to Y$ *with* $\operatorname{lip} G(\bar{x}) = 0$. *Then*

$$\operatorname{lip}(\mathbf{F} + G)^{-1}(\bar{y} + G(\bar{x})\,|\,\bar{x}) = \operatorname{lip} \mathbf{F}^{-1}(\bar{y}\,|\,\bar{x})).$$

From this result we obtain an "inverse function" version of 1.1:

COROLLARY 2.2 *Let* $F : X \rightrightarrows Y$, $f : X \to Y$ *and* $g : X \to Y$ *be such that*

$$\bar{y} \in f(\bar{x}) + F(\bar{x}), \quad g(\bar{x}) = f(\bar{x}) \quad \textit{and} \quad \operatorname{lip}(f - g)(\bar{x}) = 0.$$

*Then*

$$\operatorname{lip}(f + F)^{-1}(\bar{y}\,|\,\bar{x}) = \operatorname{lip}(g + F)^{-1}(\bar{y}\,|\,\bar{x}).$$

*Proof.* Take $\mathbf{F} = F + f$ and $G = g - f$. Then $\mathbf{F} + G = g + F$ and $\operatorname{lip} G(\bar{x}) = 0$ and we can apply 2.1. ■

The inverse function version of the original Robinson's theorem then becomes

COROLLARY 2.3 *For* $X$ *a Banach space, let* $f : X \to X^*$ *be strictly differentiable at* $\bar{x}$ *and* $C$ *be a convex and closed set. Then*

$$\operatorname{lip}(f + N_C)^{-1}(0\,|\,\bar{x}) = \operatorname{lip}(f(\bar{x}) + \nabla f(\bar{x})(\cdot - \bar{x}) + N_C(\cdot))^{-1}(0\,|\,\bar{x}).$$

We conclude this section with two remarks. First, observe that in 2.3 the mapping $\mathbf{F}$ and the function $G$ play similar (but not identical!) roles. Can we have a set-valued mapping $\mathbf{G}$ instead of the function $G$ and replace the Lipschitz modulus of $G$ by the modulus of a lipschitzian single-valued localization of $\mathbf{G}$? If yes, the claim would be more appealing esthetically. However, the answer of this question is *no, in general*, as the following example shows.

EXAMPLE 2.1 *Let* $X = Y = I\!R$ *and let* $\mathbf{F}(x) = \{2x, 1\}, \mathbf{G}(x) = \{0, -1\}$. *Both* $\mathbf{F}^{-1}$ *and* $\mathbf{G}$ *have lipschitzian localizations at* $0$ *for* $0$ *and* $\operatorname{lip} \mathbf{F}^{-1}(0\,|\,0) = .5$, $\operatorname{lip} \mathbf{G}(0\,|\,0) = 0$, *but* $(\mathbf{F}+\mathbf{G})(x) = \{2x, 2x-1, 1, 0\}$ *and* $\operatorname{lip}(\mathbf{F}+\mathbf{G})^{-1}(0\,|\,0) = \infty$.

The second remark concerns a property weaker than strong regularity, recently studied by A. Levy (2000). Up to certain notational adjustments, a mapping $A : Y \rightrightarrows X$ with $(\bar{y}, \bar{x}) \in \operatorname{gph} A$ is called *calm* at $\bar{y}$ for $\bar{x}$ in Levy (2000) when there exist a constant $c > 0$ and neighborhoods $U$ of $\bar{x}$ and $V$ of $\bar{y}$ such that the truncated mapping $V \ni y \mapsto A(y) \cap U$ is single-valued, that is, $V \ni y \mapsto x(y) := A(y) \cap U$ is a function and also $\|x(y) - \bar{x}\| \leq c\|y - \bar{y}\|$ for all $y \in V$. Our question is whether the calmness of the inverse of a mapping is preserved under perturbations by a single-valued mapping with a small Lipschitz modulus, as it is for the strong regularity. The answer turns out to be again negative.

EXAMPLE 2.2 *Consider a multivalued mapping $F$ which is a concatenation of infinitely many linear pieces symmetric with respect to the vertical axis and with lengths and slopes going to zero, namely:*

$$F(x) = \begin{cases} \frac{1}{4n+1}x - \frac{2}{4n+1} & \text{for } -\frac{1}{2n+1} \leq x \leq \frac{1}{2n} \\ -\frac{1}{4n+3}x + \frac{4n+2}{(2n+1)(4n+3)} & \text{for } -\frac{1}{2n+1} \leq x \leq \frac{1}{2n+2} \end{cases}$$

*for* $n = 1, 2, \ldots$. *The inverse of this mapping is single-valued, calm at the origin but not Lipschitz continuous near the origin. Now for every* $\varepsilon > 0$ *the mapping* $x \mapsto F(x) + \varepsilon x$ *has the property that there exists a neighborhood of* $y = 0$ *such that the inverse mapping* $(F + \varepsilon I)^{-1}$ *is not single-valued in this neighborhood. Indeed, close to the origin (where $n$ is very large) all linear pieces of* $F + \varepsilon I$ *will have positive slopes and then the inverse mapping may have two or more values for the same $y$. We can modify the example by taking pieces of parabolas with minimum on the $y$ axis, then perturb by* $\varepsilon x^2$ *and again obtain a multivalued inverse; in this case the Lipschitz modulus of the perturbation at zero is zero, that is, we can destroy the property of calmness without changing the first-order information about the mapping.*

## 3. Semidifferentiable graphical localization

In this section we consider a class of functions, called *semidifferentiable* functions, that play an important role in the analysis of variational problems with constraints. Throughout, $X$, $Y$ and $P$ are Banach spaces.

DEFINITION 3.1 *A function $f : X \to Y$ is said to be semidifferentiable at $\bar{x} \in \operatorname{dom} f$ when there is a continuous and positively homogeneous function, denoted $Df(\bar{x})$, such that*

$$f(\bar{x}+h) = f(\bar{x}) + Df(\bar{x})h + o(\|h\|).$$

By defining the norm of $Df(\bar{x})$ as

$$\|Df(\bar{x})\| = \sup_{x \in \boldsymbol{B}} \|Df(\bar{x})x\|,$$

when $Df(\bar{x})$ exists, we obtain

$$\|Df(\bar{x})\| = \limsup_{x \to \bar{x},\ x \neq \bar{x}} \frac{\|f(x) - f(\bar{x})\|}{\|x - \bar{x}\|} \leq \operatorname{lip} f(\bar{x}). \tag{10}$$

If $Df(\bar{x})$ is a linear mapping then $f$ is Fréchet differentiable at $\bar{x}$. The simplest example of a semidifferentiable function is any function $f$ for which $f(x) - f(\bar{x})$ is a positively homogeneous function; then $Df(\bar{x}) = f(x) - f(\bar{x})$. Recall that a function $\varphi$ is positively homogeneous when $0 \in \operatorname{dom} \varphi$ and $\varphi(\lambda x) = \lambda \varphi(x)$ for all $x \in \operatorname{dom} \varphi$ and $\lambda > 0$, i.e. when $\operatorname{gph} \varphi$ is a cone in $X \times Y$.

With a slight abuse of notation, for a set-valued mapping $S : X \rightrightarrows Y$ with a single-valued graphical localization $s$ around $(\bar{x}, \bar{y})$ we denote $DS(\bar{x}\,|\,\bar{y}) := Ds(\bar{x})$. The semiderivatives, also called Bouligand derivatives, have been introduced in various ways in the literature, for a thorough treatment see Malanowski (2001).

Our first result is in line of 2.1.

THEOREM 3.1 *Consider a mapping $\mathbf{M} : Y \rightrightarrows X$ with $\operatorname{lip} \mathbf{M}(\bar{y}\,|\,\bar{x}) < \infty$ and suppose that its graphical localization $m(\cdot)$ around $(\bar{y}, \bar{x})$ is semidifferentiable at $\bar{y}$. Consider also a function $h : P \times X \to Y$ which satisfies $\operatorname{lip} h(\bar{p}, \bar{x}) < \infty$ and $h(\cdot, \bar{x})$ is semidifferentiable at $\bar{p}$ with semiderivative $D_p h(\bar{p}, \bar{x})$. In addition, assume that*

$$\operatorname{lip} \mathbf{M}(\bar{y}\,|\,\bar{x}) \cdot \operatorname{lip}_x h(\bar{p}, \bar{x}) = 0. \tag{11}$$

*Then the mapping*

$$P \ni p \mapsto \mathbf{N}(p) := \left\{ x \in X \;\middle|\; x \in \mathbf{M}(h(p,x)) \right\}$$

*has a lipschitzian single-valued graphical localization around $(\bar{p}, \bar{x})$ which is semidifferentiable at $\bar{p}$ and its semiderivative is*

$$D\mathbf{N}(\bar{p}\,|\,\bar{x}) = D\mathbf{M}(\bar{y}\,|\,\bar{x}) \circ D_p h(\bar{p}, \bar{x}).$$

*Proof.* The existence of a lipschitzian single-valued graphical localization $x(\cdot)$ of $\mathbf{N}$ around $(\bar{p}, \bar{x})$ follows from 2.1. Indeed, the condition $\operatorname{lip} h(\bar{p}, \bar{x}) < \infty$ implies that $h$ is Lipschitz continuous with respect to $p$ near $\bar{p}$ uniformly in $x$ near $\bar{x}$

and then the Lipschitz continuity of the single-valued localization follows from (8).

Let $\varepsilon > 0$ and let $Q$ be a neighborhood of $\bar{p}$ such that, for a fixed $p \in Q$, the function $x \mapsto m(h(p,x))$ is Lipschitz continuous with a constant $\varepsilon$. That such choice of $Q$ is possible follows from the condition (11) and from the continuity of $x(\cdot)$ and $h(\cdot)$ around $\bar{p}$ and $(\bar{p},\bar{x})$, respectively. Indeed, taking $Q$ smaller if necessary, for any arbitrary small neighborhood $U$ of $\bar{x}$ and any arbitrary small neighborhood $V$ of $\bar{y}$ one can have $x(p) \in U$ whenever $p \in Q$ and $h(p,x) \in V$ whenever $p \in Q$ and $x \in U$. Further, from (11), for some $\alpha > \operatorname{lip} \mathbf{M}(\bar{y}\,|\,\bar{x}) = \operatorname{lip} m(\bar{y})$ and $\beta > \operatorname{lip}_x h(\bar{p},\bar{x})$ such that $\alpha\beta \le \varepsilon$, we take $U$, $V$ and $Q$ small enough such that $m(\cdot)$ is Lipschitz in $V$ with a constant $\alpha$ and $h(p,\cdot)$ is Lipschitz (with respect to $x$) in $U$ uniformly in $p \in Q$ with a constant $\beta$. We can also take $Q$ and $V$ even smaller, if needed, such that for $p \in Q$,

$$\|h(p,\bar{x}) - h(\bar{p},\bar{x}) - D_p h(\bar{p},\bar{x})(p-\bar{p})\| \le \varepsilon \|p - \bar{p}\| \tag{12}$$

and for $y \in V$,

$$\|m(y) - m(\bar{y}) - D\mathbf{M}(\bar{y}\,|\,\bar{x})(y-\bar{y})\| \le \varepsilon \|y - \bar{y}\|. \tag{13}$$

For $p \in Q$ we write

$$\begin{aligned}
&\|x(p) - \bar{x} - D\mathbf{M}(\bar{y}\,|\,\bar{x})(D_p h(\bar{p},\bar{x})(p-\bar{p}))\| \\
&= \|m(h(p,x(p))) - m(\bar{y}) - D\mathbf{M}(\bar{y}\,|\,\bar{x})(D_p h(\bar{p},\bar{x})(p-\bar{p}))\| \\
&\le \|m(h(p,x(p))) - m(h(p,\bar{x}))\| \\
&+ \|m(h(p,\bar{x})) - m(\bar{y} + D_p h(\bar{p},\bar{x})(p-\bar{p}))\| \\
&+ \|m(\bar{y} + D_p h(\bar{p},\bar{x})(p-\bar{p})) - m(\bar{y}) - D\mathbf{M}(\bar{y}\,|\,\bar{x})(D_p h(\bar{p},\bar{x})(p-\bar{p}))\|.
\end{aligned} \tag{14}$$

The expressions in the right-hand side of (14) are now estimated. For the first expression we use the assumption (11), that is, with the so-chosen Lipschitz constants $\alpha$ of $m(\cdot)$ and $\beta$ of $h(p,\cdot)$:

$$\begin{aligned}
\|m(h(p,x(p))) - m(h(p,\bar{x}))\| &\le \alpha \|h(p,x(p)) - h(p,\bar{x})\| \\
&\le \alpha\beta \|p-\bar{p}\| \le \varepsilon \|p - \bar{p}\|.
\end{aligned} \tag{15}$$

The second expression is estimated by using the Lipschitz constant of $m(\cdot)$ and (12):

$$\begin{aligned}
&\|m(h(p,\bar{x})) - m(\bar{y} + D_p h(\bar{p},\bar{x})(p-\bar{p}))\| \\
&\quad \le \alpha \|h(p,\bar{x}) - h(\bar{p},\bar{x}) - D_p h(\bar{p},\bar{x})(p-\bar{p})\| \le \alpha\varepsilon \|p-\bar{p}\|.
\end{aligned} \tag{16}$$

From (13), for the third expression we obtain:

$$\begin{aligned}
&\|m(\bar{y} + D_p h(\bar{p},\bar{x})(p-\bar{p})) - m(\bar{y}) - D\mathbf{M}(\bar{y}\,|\,\bar{x})(D_p h(\bar{p},\bar{x})(p-\bar{p}))\| \\
&\quad \le \varepsilon \|D_p h(\bar{p},\bar{x})(p-\bar{p}))\| \le \varepsilon \|D_p h(\bar{p},\bar{x})\| \|p-\bar{p}\|.
\end{aligned} \tag{17}$$

Using (15), (16) and (17) in (14) and taking into account (10) we obtain

$$\|x(p) - \bar{x} - D\mathbf{M}(\bar{y}\,|\,\bar{x})(D_p h(\bar{p},\bar{x})(p-\bar{p}))\| \le \varepsilon(1 + \alpha + \operatorname{lip}_x h(\bar{p},\bar{x}))\|p-\bar{p}\|.$$

Since $\varepsilon$ can be arbitrary small, the last expression is $o(\|p-\bar{p}\|)$ and thus the theorem is proved. ■

Theorem 3.2 *Given* $F : X \rightrightarrows Y$ *and* $f : P \times X \to Y$ *with* $\bar{y} \in f(\bar{p}, \bar{x}) + F(\bar{x})$, *let* $\operatorname{lip} f(\bar{p}, \bar{x}) < \infty$ *and* $f(\cdot, \bar{x})$ *be semidifferentiable at* $\bar{p}$. *Let* $g : X \to Y$ *be a first-order approximation of* $f$ *at* $(\bar{p}, \bar{x})$ *and let* $(g + F)^{-1}$ *have a lipschitzian single-valued localization around* $(\bar{y}, \bar{x})$ *which is semidifferentiable at* $\bar{y}$. *Then the mapping*

$$p \mapsto S(p) := \{x \in X \mid f(p, x) + F(x) \ni 0\} \tag{18}$$

*has a lipschitzian single-valued localization around* $(\bar{p}, \bar{x})$ *which is semidifferentiable at* $\bar{p}$ *and moreover*

$$DS(\bar{p} | \bar{x}) = D(g + F)^{-1}(\bar{y} | \bar{x}) \circ (-D_p f(\bar{p}, \bar{x})).$$

*Proof.* The existence of a lipschitzian single-valued localization follows from 2.2. Further, for $\mathbf{M} = (g + F)^{-1}$ and $h(p, x) = g(x) - f(p, x)$ by the first-order approximation we have that $\operatorname{lip}_x h(\bar{p}, \bar{x}) = 0$ and $h(\cdot, \bar{x})$ is semidifferentiable at $\bar{p}$. Clearly, $\operatorname{lip} h(\bar{p}, \bar{x}) \leq \operatorname{lip} g(\bar{x}) + \operatorname{lip} f(\bar{p}, \bar{x}) < \infty$ and then the claim follows from 3.1. ■

The inverse mapping form of the above theorem is

Theorem 3.3 *Consider a set-valued mapping* $\mathbf{F} : X \rightrightarrows Y$ *with* $\operatorname{lip} \mathbf{F}^{-1}(\bar{y} | \bar{x}) < \infty$ *and let the associated single-valued localization of* $\mathbf{F}^{-1}$ *around* $(\bar{y}, \bar{x})$ *be semidiferentiable at* $\bar{y}$. *Consider also a function* $G : X \to Y$ *with* $\operatorname{lip} G(\bar{x}) = 0$. *Then*

$$\operatorname{lip}(\mathbf{F} + G)^{-1}(\bar{y} + G(\bar{x}) | \bar{x}) = \operatorname{lip} \mathbf{F}^{-1}(\bar{y} | \bar{x})$$

*and the associated single-valued graphical localization of* $(\mathbf{F} + G)^{-1}$ *around* $(\bar{y} + G(\bar{x}), \bar{x})$ *is semidiferentiable at* $\bar{y} + G(\bar{x})$.

The results in this section partially strengthen author's results from Dontchev (1995), where it is also shown that, with appropriate modifications, the semidifferentiability property in the above analysis can be replaced without major changes by other types of differentiability, such as directional differentiability or Fréchet differentiability. An application of these results to optimal control is given in Malanowski (2001b).

## References

Bonnans, J. F., and Shapiro, A. (2000) *Perturbation Analysis of Optimization Problems. Springer Series in Operations Research.* Springer-Verlag, New York.

Dontchev, A.L. (1995) Implicit function theorems for generalized equations. *Math. Programming*, **70**, Ser.A, 91–106.

Klatte, D. and Kummer, B. (2002) Nonsmooth equations in optimization. Regularity, calculus, methods and applications. *Nonconvex Optimization and its Applications*, **60**, Kluwer Academic Publishers, Dordrecht.

Levy A.B. (2000) Calm minima in parameterized finite-dimensional optimization, *SIAM J. Optim.* **11**, 160–178.

Malanowski, K. (2001a) Stability and sensitivity analysis for optimal control problems with control-state constraints. *Dissertationes Math.* **394.**

Malanowski, K. (2001b)Bouligand differentiability of solutions to parametric optimal control problems. *Numer. Funct. Anal. Optim.* **22**, 973–990.

Robinson, S.M. (1980)Strongly regular generalized equations, *Math. of Oper. Research* **5**, 43–62.

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

Wets, R.J.-B. and Rockafellar, R.T. (1997) Variational Analysis, Springer-Verlag, Berlin.