Ann. Inst. Henri Poincaré,

Vol. 7, n° 3, 1990, p. 183-234. *Analyse non linéaire*

Some inverse mapping theorems

by

Hélène FRANKOWSKA

Université de Paris-Dauphine,
Place du Maréchal-De-Lattre-de-Tassigny,
75775, Paris Cedex 16, France

ABSTRACT. – We prove several first and high order inverse mapping theorems for set-valued maps from a complete metric space to a Banach space and study the stability of the open mapping principle. The obtained results allow to investigate questions of controllability of finite and infinite dimensional control systems, necessary conditions for optimality, implicit function theorem, stability of constraints with respect to a parameter. Applications to problems of optimization, control theory and nonsmooth analysis are provided.

Key words : Controllability, high order inverse, implicit function theorem, inverse mapping theorem, multiplier rule, nonsmooth analysis, open mapping principle, optimality, reachability, stability, variation of set-valued map.

RÉSUMÉ. – On démontre des théorèmes d'inversion locale d'ordre quelconque pour des correspondances définies sur un espace métrique complet. Les résultats sont appliqués à quelques problèmes des contrôles.

Mots clés : Inverse mapping, reachability, controllability, optimal control.

Classification A.M.S. : 26 E 25, 34 A 60, 49 B 10, 49 E 05, 49 E 15, 49 E 30, 93 B 05, 93 C 15, 93 C 20.

Annales de l'Institut Henri Poincaré - Analyse non linéaire - 0294-1449
Vol. 7/90/03/183/52/$7.20/

1. INTRODUCTION

Consider a set-valued map $G:X \rightarrow Y$ between two metric spaces and define its inverse $G^{-1}:Y \rightarrow X$ by

$$G^{-1}(y)=\{x\in X \mid y\in G(x)\}$$

Let $\bar{x}\in X$, $\bar{y}\in G(\bar{x})$ and let $B_h(\bar{x})$ denote the closed ball in X centered at $\bar{x}$ with radius $h>0$. This paper is mainly concerned with sufficient conditions for:

1. the uniform open mapping principle at $(\bar{x}, \bar{y})$: $\exists\, \varepsilon>0$, $k>0$, $\rho>0$ such that

$$\forall (x, y)\in \text{Graph}(G) \cap B_\varepsilon(\bar{x}) \times B_\varepsilon(\bar{y}),$$
$$\forall h\in[0, \varepsilon], \qquad B_{\rho h^k}(y)\in G(B_h(x))$$

2. the Hölder regularity of the inverse map G^{-1} at $(\bar{y}, \bar{x})$: $\exists\, k>0$, $\varepsilon>0$, $L>0$ such that

$$\forall (x_1, y_1)\in \text{Graph}(G) \cap B_\varepsilon(\bar{x}) \times B_\varepsilon(\bar{y}),$$
$$\forall y_2\in B_\varepsilon(\bar{y}), \qquad \text{dist}(x_1, G^{-1}(y_2))\leqq L\, d_Y(y_1, y_2)^{1/k}$$

3. necessary conditions verified by boundary points of $G(B_\varepsilon(\bar{x}))$.

A classical result of functional analysis states that if a (single-valued) C^1-function $f:X \rightarrow Y$ between two Banach spaces has a surjective derivative $f'(\bar{x})$ at a point $\bar{x}\in X$, then for all $h>0$, $f(\bar{x})\in \text{Int} f(B_h(\bar{x}))$ (*i.e.* the open principle holds true) and the set-valued map f^{-1} is roughly speaking Lipschitzian at $f(\bar{x})$. It implies in particular that if $f'(\bar{x})$ is surjective, then $\text{Ker} f'(\bar{x})$ is tangent to the level set $\{x\in X: f(x)=f(\bar{x})\}$ [31]. We refer to [11], [10], [28] for historical comments and an extensive bibliography. The open mapping principle part of the above theorem is sometimes referred as Graves theorem ([24], [25]).

However the above classical result is not strong enough to answer many questions arising in Control Theory and Optimization:

We may have to deal with maps defined on metric spaces (which have no much regularity) rather than with C^1, single-valued functions. On the other hand the sets of constraints are given by set-valued maps. This is then a first source of motivations to use new tools adapted to these purposes.

Actually set-valued maps are in the background, even when many efforts where devoted to hide them. Indeed there have been many attempts to overcome the difficulties, and most of them are actually based on a careful construction of a selection $f(x)\in G(x)$ to which one or another open mapping result can be applied. However very often this is neither a direct nor a simple way to follow.

Let us mention also that beside the above theorem used together with such selection f [6], several different open mapping arguments have been applied to f, for instance the one based on Brouwer's fixed point theorem (*see* for example proof of Pontriagin' principle in [30]), a "degree theory" open mapping result [39], extensions of Grave's theorem to nonsmooth functions [43], etc. We also refer to [33], where a number of fixed point theorems and their applications to infinite dimensional control problems are given.

Our strategy is then to deal directly with inverse and open mapping theorems for set-valued maps from a complete metric space to a Banach space and to replace the notion of derivative (which needs a linear structure) by a notion of variation, which describes the infinitesimal behavior of a map at a given point.

The inverse function theorem for set-valued maps is also a very convenient tool to be applied to optimization problems, even those whose data are given by single valued functions. We refer to [1], [4], [28] and their references, where this issue is illustrated by many examples.

Most extensions of the above classical result are of the first order involving surjectivity of the first derivative. Such assumption excludes from consideration those functions whose derivative is not surjective or simply vanishes. A high order open mapping principle for single valued maps taking their values in a finite dimensional space was proved in [25] using Brouwer's fixed point theorem and for set-valued maps in [18] on the basis of Ekeland's principle. I would like also to acknowledge the private communication of J. Borwein, that, still using Ekeland's principle, it is possible to show that the sufficient condition for openness proved in [18] implies as well the Hölder continuity of the inverse.

Finally a convenient extension of the inverse mapping theorem is necessary to treat nonsmooth problems. Some generalizations in this direction can be found in [8], [43].

In this paper we prove first and high order sufficient conditions for openness and regularity of the inverse map which allows to obtain several new results concerning controllability, optimality and Lipschitzian realizations. Sufficient conditions for invertibility are expressed in terms of *variations* of set-valued maps defined on metric spaces.

Variations measure infinitesimal behavior of a map and seem to be a very (may be the most) natural notion to be used to study the uniform open mapping principle (*see* Remark following Definition 5.1 from Section 5). In several examples of applications provided here, we show how variations can be computed. In particular this leads to a short and direct proof of the maximum principle in control theory (both for finite and infinite dimensional cases).

In summary our extensions deal with:

1. *nonsmooth* functions and *set-valued maps*;
2. set-valued maps defined on a *complete metric space*;

3. maps taking their values in a *Banach space*;
4. *high order* sufficient conditions.

The strength of these results is showed by examples of applications to:
1. reachability of nonlinear infinite dimensional control systems;
2. necessary conditions for optimality;
3. small time local controllability;
4. some problems of nonsmooth analysis.

The outline of the paper is as follows. We first show equivalence of the uniform open mapping principle and Hölder continuity of the inverse (Section 2). So the problem of regular inverse reduces to sufficient conditions for the uniform open mapping principle. We use Ekeland's variational principle to investigate this problem. In particular we prove that whenever Y is a smooth Banach space, then the uniform open mapping principle is equivalent to a "convex uniform open mapping principle", which can then be applied together with separation theorems (Section 3). In [17] we derived from these results some second order conditions for invertibility of a C^2-function taking its values in a Hilbert space and for stability with respect to a parameter of a system defined by inequality constraints.

Applications of Ekeland's principle appear often in nonsmooth analysis to derive necessary conditions for optimality and sufficient conditions for invertibility. In this paper Ekeland's principle is rather used to prove stability of the uniform open mapping principle (Section 4). Necessary conditions for optimality can be seen as a violation of the open mapping principle. For nonsmooth problems, necessary conditions can be seen as a violation of the uniform open mapping principle for smooth approximations. This approach to nonsmooth problems was pioneered by J. Warga ([41]-[43]). In Section 13, we derive several "nonsmooth" results based on stability of the open mapping principle.

To express sufficient conditions for the uniform open mapping principle we introduce in Section 5 "variations" of set-valued maps, which describe infinitesimal changes of a map on a neighborhood of a given point. We shall consider two different types of variations. The first one is called contingent variation and is related to the (first) Gâteaux derivative. It allows to prove the first order sufficient conditions. The second type of variations is much more regular (similar to continuous Fréchet derivative) and is defined for all orders. Naturally it leads to higher order results.

Several extensions of the first order conditions to set-valued maps on Banach spaces can be found in [1], [4], [28]. A high order open mapping principle via high order variations of set-valued maps was proposed in [18]. Their proofs use Ekeland's principle but now can be derived directly from Theorem 2.2 (Section 2). The proof of Theorem 2.2 is based on a constructive argument, similar to the one used in [24], [11], which allows to estimate the Hölder constant, the neighborhood where the map can be inverted and errors. Although several proofs of this paper are also based on the variational principle, it is applied mainly to bring some convexity arguments, while in earlier works this was not exploit at all.

The first order inverse function theorems for a map G taking its values in a Banach space Y are proved in Section 6. Special attention is given to the case when the norm of Y is Gâteaux differentiable away from zero. This also allows to derive necessary conditions satisfied by boundary points of the image of G. They are used to prove necessary conditions for optimality for an abstract infinite dimensional mathematical programming problem in Section 11, which in turn is applied to an infinite dimensional optimal control problem with end point constraints and to a semilinear optimal control problem.

First order inverse function theorems are also used to investigate Lipschitz behavior of controls for finite and infinite dimensional control systems (Section 10) and the implicit function theorem (Section 9). This last theorem allows to make a Lipschitz realization of an implicit dynamical system.

A high order inverse function theorem for maps with values in a uniformly smooth Banach space is proved in Section 7. An application of this result to question of small time local controllability is given in Section 12. Finally Section 13 is devoted to some finite dimensional nonsmooth problems. One can find in [17] some further applications of set-valued inverse mapping theorems.

CONTENTS

2. UNIFORM OPEN MAPPING PRINCIPLE AND INVERSE MAPPING THEOREM

Consider a set-valued map G from a complete metric space (X, d_X) to a metric space (Y, d_Y). That is for every $x \in X$, $G(x)$ is a (possibly empty) subset of Y. Recall that graph of G is a subset of the product space $X \times Y$:

$$\text{Graph}(G) = \{(x, y) \mid y \in G(x)\}.$$

When it is not otherwise specified explicitly, we shall use on it the following metric:

$$\forall (x, y), \quad (x', y') \in \text{Graph}(G), \qquad d((x, y), (x', y')) = d_X(x, x') + d_Y(y, y').$$

In this paper we restrict our attention to those maps whose graph is closed. From now on we posit such an assumption. The inverse map G^{-1} is defined by

$$\forall y \in Y, \quad G^{-1}(y) := \{x \in X \mid y \in G(x)\}.$$

In the other words $(x, y) \in \text{Graph}(G)$ if and only if $(y, x) \in \text{Graph}(G^{-1})$.

In this section we study a relationship between the uniform open mapping principle and the regularity of the inverse map G^{-1}. For all $x \in X$, $h > 0$ denote by $\mathring{B}_h(x)$ [respectively $B_h(x)$] the open (respectively closed) ball in X of center x and radius $h > 0$.

THEOREM 2.1. — *Let* $k > 0$, $\rho > 0$, $x_0 \in X$, $y_0 \in G(x_0)$. *The following statements are equivalent:*

(i) *For all* $(x, y) \in \text{Graph}(G)$ *near* (x_0, y_0) *and all small* $h > 0$

$$\mathring{B}_{\rho h^k}(y) \subset G(B_h(x)). \tag{1}$$

(ii) *For all* $(x_1, y_1) \in \text{Graph}(G)$ *near* (x_0, y_0) *and all* $y_2 \in Y$ *near* y_0

$$\text{dist}_X(x_1, G^{-1}(y_2)) \leq \rho^{-1/k} d_Y(y_1, y_2)^{1/k}. \tag{2}$$

Remark. — When Y is a Banach space and $\mathring{B}$ denotes the open unit ball in Y, then the inclusion (1) can be formulated as

$$y + \rho h^k \mathring{B} \subset G(B_h(x))$$

or equivalently as

$$\rho \mathring{B} \subset \frac{G(B_h(x)) - y}{h^k}.$$

The implication (ii) ⇒ (i) in the above theorem is immediate. The opposite claim results from the more precise:

THEOREM 2.2. – *Consider* $y_0 \in G(x_0)$. *If there exist* $k>0$, $\varepsilon>0$, $\rho>0$, $0<\beta<1$ *such that for all* $(x, y) \in \text{Graph}(G) \cap B_\varepsilon(x_0) \times B_\varepsilon(y_0)$ *and all* $h \in [0, \varepsilon]$

$$\sup_{b \in B_{\rho h^k}(y)} \text{dist}_Y(b, G(B_h(x))) \leq \beta\rho h^k$$

then, for all $(x_1, y_1) \in \text{Graph}(G) \cap B_{\varepsilon/2}(x_0) \times B_{\varepsilon/2}(y_0)$, *for all* $h \geq 0$ *satisfying* $\max\left\{\frac{h}{1-\beta^{1/k}}, 2\rho h^k\right\} < \frac{\varepsilon}{2}$ and all $y_2 \in B_{\rho h^k}(y_1)$

$$\text{dist}_X(x_1, G^{-1}(y_2)) \leq \frac{1}{1-\beta^{1/k}} h$$

or equivalently for all $(x_1, y_1) \in \text{Graph}(G) \cap B_{\varepsilon/2}(x_0) \times B_{\varepsilon/2}(y_0)$ *and all* $y_2 \in Y$ *satisfying* $d_Y(y_1, y_2) < \min\left\{\frac{\varepsilon}{4}, \rho\left(\frac{\varepsilon}{2}\right)^k \left(1-\beta^{1/k}\right)^k\right\}$

$$\text{dist}_X(x_1, G^{-1}(y_2)) \leq \frac{1}{1-\beta^{1/k}} \frac{1}{\rho^{1/k}} d_Y(y_1, y_2)^{1/k}.$$

Proof. – Fix any $\beta<\alpha<1$ such that $h/(1-\alpha^{1/k}) \leq \frac{\varepsilon}{2}$. Then for all $(x, y) \in \text{Graph}(G) \cap B_\varepsilon(x_0) \times B_\varepsilon(y_0)$ and all $h \in]0, \varepsilon]$

$$\sup_{b \in B_{\rho h^k}(y)} \text{dist}_Y(b, G(B_h(x))) < \alpha\rho h^k. \tag{3}$$

Let x_1, y_1, y_2, h as in the conclusion of theorem. It is enough to consider the case $h \neq 0$. We look for $x_2 \in G^{-1}(y_2)$ satisfying $d_X(x_1, x_2) \leq h/(1-\alpha^{1/k})$ as the limit of a sequence we shall built. Set $u_0 = x_1$. By (3) there exists $(u_1, v_1) \in \text{Graph}(G)$ such that $d_X(u_0, u_1) = d_X(x_1, u_1) \leq h$, $d_Y(v_1, y_2) < \alpha\rho h^k$. Assume that we already constructed $(u_i, v_i) \in \text{Graph}(G)$, $i = 1, \ldots, n$ such that

$$d_X(u_{i-1}, u_i) \leq \alpha^{(i-1)/k} h \tag{4}$$

$$d_Y(v_i, y_2) < \rho\alpha^i h^k = \rho(\alpha^{i/k} h)^k. \tag{5}$$

Then

$$d_X(x_1, u_i) \leq \sum_{j=1}^{i} d_X(u_{j-1}, u_j) \leq h \sum_{j=0}^{i-1} \alpha^{j/k} \leq \frac{h}{1-\alpha^{j/k}} \tag{6}$$

and

$$d_X(x_0, u_i) \leqq d_X(x_0, x_1) + d_X(x_1, u_i) \leqq \frac{\varepsilon}{2} + \frac{h}{1-\alpha^{1/k}} \leqq \varepsilon$$

$$d_Y(y_0, v_i) \leqq d_Y(y_0, y_1) + d_Y(y_1, y_2) + d_Y(y_2, v_i) \leqq \frac{\varepsilon}{2} + \rho h^k + \rho \alpha^i h^k \leqq \varepsilon.$$

Hence by (3) and by (5) applied to (u_n, v_n), there exists $(u_{n+1}, v_{n+1}) \in \text{Graph}(G)$ such that

$$d_X(u_n, u_{n+1}) \leqq \alpha^{n/k} h, \qquad d_Y(v_{n+1}, y_2) < \rho \alpha^{n+1} h^k.$$

Observe that (4) implies that $\{u_i\}$ is a Cauchy sequence and that (5) implies that $\lim_{i \to \infty} v_i = y_2$. Let x_2 be the limit of $\{u_i\}$. Since Graph (G) is closed, $(x_2, y_2) \in \text{Graph}(G)$ and thus $x_2 \in G^{-1}(y_2)$. Moreover by (6), $d_X(x_1, x_2) \leqq \frac{h}{1-\alpha^{1/k}}$ and therefore

$$d_X(x_1, G^{-1}(y_2)) \leqq \frac{1}{1-\alpha^{1/k}} h.$$

Since $\alpha \in]\beta, 1[$ may be chosen arbitrary close to β, the proof is complete.

3. UNIFORM OPEN MAPPING PRINCIPLE IN SMOOTH BANACH SPACES

In the previous section we have shown that the uniform open mapping principle (1) is a necessary and sufficient condition for the "Hölder continuity of the inverse map" (2). However verification of the open mapping principle may be a difficult task. In this section we replace it by a "convex uniform open mapping principle" which, thanks to separation theorems, is more simple to deal with. We assume that Y is a Banach space and that its norm $\|\cdot\|$ is Fréchet differentiable away from zero. We denote by d the metric of the complete metric space X and by $\overline{co}$ the closed convex hull. We start by a first order result.

THEOREM 3.1. – *Let* $y_0 \in G(x_0)$. *If for some* $\rho > 0$, $\varepsilon > 0$, $M > 0$ *and all* $(x, y) \in \text{Graph}(G) \cap B_\varepsilon(x_0) \times B_\varepsilon(y_0)$ *and all* $h \in]0, \varepsilon]$

$$\rho \mathring{B} \subset \overline{co}\left(\frac{G(B_h(x)) - y}{h} \cap MB\right) \tag{7}$$

or equivalently

$$\inf_{\|p\|_{Y^*}=1} \sup\left\{\langle p, v\rangle \,\middle|\, v \in \frac{G(B_h(x)) - y}{h} \cap MB\right\} \geqq \rho$$

then for all $(x, y) \in \text{Graph}(G) \cap B_{\varepsilon/2}(x_0) \times B_{\varepsilon/2}(y_0)$ *and all* $h \in]0, \frac{\varepsilon}{2}[$

$$\rho \mathring{B} \subset \frac{G(B_h(x)) - y}{h}. \tag{8}$$

Remark

(*a*) Assumption (7) may be seen as the convex uniform open mapping principle (compare with Remark following Theorem 2.1).

(*b*) It is clear that if (8) holds true on a neighborhood of $(x_0, y_0, 0)$ in $\text{Graph}(G) \times R_+$ then so does (7). Hence uniform open mapping principle and the convex uniform open mapping principle are equivalent in those spaces whose norm is Fréchet differentiable away from zero. □

Proof. – It is enough to prove that for every $\lambda > 0$ and all x, y, h as in the conclusion of the theorem

$$\frac{\rho}{1+\lambda} \mathring{B} \subset \frac{G(B_h(x)) - y}{h}. \tag{9}$$

Fix $\lambda > 0$ and assume for a moment that for some $(t, z) \in \text{Graph}(G) \cap B_{\varepsilon/2}(x_0) \times B_{\varepsilon/2}(y_0)$ and $\bar{h} \in]0, \frac{\varepsilon}{2}[$, there exists

$$\bar{y} \in z + \frac{\rho \bar{h}}{1+\lambda} \mathring{B}, \qquad \bar{y} \notin G(B_{\bar{h}}(t)) \tag{10}$$

Define $0 < \Theta < 1$ by $\Theta^2 = \| z - \bar{y} \| (1+\lambda)/\rho \bar{h}$. Applying the Ekeland variational principle [12], [13] to the complete metric space $K := \text{Graph}(G) \cap B_{\bar{h}}(t) \times Y$ with the metric

$$d_{X \times Y}((x, y), (x', y')) = d(x, x') + \frac{\lambda}{M} \| y - y' \| \tag{11}$$

and the continuous function $(x, y) \to \| y - \bar{y} \|$ we prove the existence of $(x, y) \in K \cap B_{\Theta \bar{h}}(t) \times B_{\Theta \bar{h}}(z)$ such that

$$\forall (u, w) \in K, \quad \| y - \bar{y} \| \leq \| w - \bar{y} \| + \frac{\Theta \rho}{1+\lambda} \left(d(u, x) + \frac{\lambda}{M} \| w - y \| \right) \tag{12}$$

From (10) we know that $y \neq \bar{y}$. By differentiability of the norm of Y, for some $p \in Y^*$ of $\| p \|_{Y^*} = 1$ and all $v \in Y$ of $\| v \| \leq M$, $\| y + hv - \bar{y} \| \leq \| y - \bar{y} \| + \langle p, hv \rangle + o(h)$, where $\lim_{h \to 0+} \frac{o(h)}{h} = 0$. Hence for all $(u, y + hv) \in K$ with $\| v \| \leq M$

$$\| y - \bar{y} \| \leq \| y - \bar{y} \| + \langle p, hv \rangle + o(h) + \frac{\Theta \rho}{1+\lambda} (d(u, x) + \lambda h)$$

and therefore, for all small $h>0$

$$\forall v \in \frac{G(B_h(x))-y}{h} \cap MB, \quad \langle p, v\rangle \geqq -\left(\Theta\rho + \frac{o(h)}{h}\right).$$

This yields that for some $\varepsilon_h \to 0+$

$$\forall v \in \overline{co}\left(\frac{G(B_h(x))-y}{h} \cap MB\right), \quad \langle p, v\rangle \geqq -\Theta\rho - \varepsilon_h.$$

Let $v \in Y$, $\|v\|<1$ be such that $\langle p, v\rangle < -\sqrt{\Theta}$. By our assumptions, for all $h \in]0, \varepsilon]$, ρv belongs to the right-hand side of (7). Hence for all small $h>0$ we have $-\sqrt{\Theta}\rho \geqq -\Theta\rho - \varepsilon_h$, which leads to a contradiction and ends the proof. □

In the high order result stated below we assume that Y is a uniformly smooth Banach space, *i.e.* its norm $\|\,.\,\|$ is uniformly Fréchet differentiable away from zero. That is, for a function $o: \mathbf{R}_+ \to \mathbf{R}_+$ satisfying $\lim_{t\to 0+} o(t)/t = 0$ and for $y \in Y$ with $0 < \|y\| \leqq 1$ there exists $J(y) \in Y^*$, $\|J(y)\|_{Y^*} = 1$ such that for every $t \in \mathbf{R}$

$$\sup_{\|v\|\leqq 1} \left|\,\|y+tv\| - \|y\| - t\langle J(y), v\rangle\right| \leqq o(|t|)$$

where $\langle .\,, .\rangle$ states for the duality pairing on $Y^* \times Y$. Recall that a uniformly smooth Banach space is reflexive. Every Hilbert space is uniformly smooth and a space Y is uniformly smooth if and only if its dual Y^* is uniformly convex. In particular L^p spaces are uniformly smooth for $1<p<\infty$ [5].

THEOREM 3.2. – *Assume that* Y *is uniformly smooth. Let* $y_0 \in G(x_0)$, $k \geqq 1$. *The following statements are equivalent:*

(i) *For some* $\rho>0$ *and for all* $(x, y) \in \operatorname{Graph}(G)$ *near* (x_0, y_0) *and all small* $h>0$

$$\rho \mathring{B} \subset \frac{G(B_h(x))-y}{h^k}.$$

(ii) *For some* $\rho>0$, $M>0$ *and for all* $(x, y) \in \operatorname{Graph}(G)$ *near* (x_0, y_0) *and all small* $h>0$

$$\rho \mathring{B} \subset \overline{co}\left(\frac{G(B_h(x))-y}{h^k} \cap MB\right).$$

(iii) *For some* $\rho>0$, $M>0$ *and for all* $(x, y) \in \operatorname{Graph}(G)$ *near* (x_0, y_0) *and all small* $h>0$

$$\inf_{\|p\|_{Y^*}=1} \sup\left\{\langle p, v\rangle \mid v \in \frac{G(B_h(x))-y}{h^k} \cap MB\right\} \geqq \rho.$$

Prof. – Because of separation theorems (ii) is equivalent to (iii). Obviously (i) yields (ii) with $M=\rho$. To prove the implication (ii) $\Rightarrow$ (i) we proceed by a contradiction argument. Assume contrary that for some $(t_i, z_i)\in \mathrm{Graph}(G)$ converging to (x_0, y_0) and $h_i \to 0+$ there exist

$$\bar{y}_i \in z_i + 2^{-k_i-2k} h_i^k \mathring{B}, \qquad \bar{y}_i \notin G(B_{h_i}(t_i)) \tag{13}$$

Applying the Ekeland variational principle [12], [13] to the complete metric space $K_i := \mathrm{Graph}(G) \cap B_{h_i}(t_i) \times Y$ and the continuous function $(x, y) \to \|y-\bar{y}_i\|^{1/k}$, we prove the existence of $(x_i, y_i)\in K_i$ such that

$$d(x_i, t_i) + \|y_i - z_i\| \leqq \frac{1}{2i} h_i, \qquad \|y_i - \bar{y}_i\|^{1/k} \leqq \frac{1}{2i^2} h_i \tag{14}$$

$$\forall (x, y) \in K_i, \qquad \|y_i - \bar{y}_i\|^{1/k} \leqq \|y-\bar{y}_i\|^{1/k} + \frac{1}{i}(d(x, x_i) + \|y-y_i\|) \tag{15}$$

From (13) we know that $y_i \neq \bar{y}_i$. Since Y is uniformly smooth for some $p_i \in Y^*$ of $\|p_i\|_{Y^*}=1$ and all y,

$$\|y-\bar{y}_i\| \leqq \|y_i - \bar{y}_i\| + \langle p_i, y-y_i\rangle + o(\|y-y_i\|).$$

Hence

$$\|y-\bar{y}_i\|^{1/k} \leqq \|y_i-\bar{y}_i\|^{1/k}\left(1+\left\langle p_i, \frac{y-y_i}{\|y_i-\bar{y}_i\|}\right\rangle + \frac{o(\|y-y_i\|)}{\|y_i-\bar{y}_i\|}\right)^{1/k}$$
$$\leqq \|y_i-\bar{y}_i\|^{1/k}\left(1+\frac{1}{k}\left\langle p_i, \frac{y-y_i}{\|y_i-\bar{y}_i\|}\right\rangle + \bar{o}\left(\frac{\|y-y_i\|}{\|y_i-\bar{y}_i\|}\right)\right)$$

for a function $\bar{o}: R_+ \to R_+$ satisfying $\lim_{h\to 0+} \frac{\bar{o}(h)}{h}=0$ and, by (15), for all $(x, y)\in K_i$

$$\|y_i-\bar{y}_i\|^{1/k} \leqq \|y_i-\bar{y}_i\|^{1/k} + \frac{1}{k}\|y_i-\bar{y}_i\|^{(1/k)-1}\langle p_i, y-y_i\rangle$$
$$+\|y_i-\bar{y}_i\|^{1/k}\bar{o}\left(\frac{\|y-\bar{y}_i\|}{\|y_i-\bar{y}_i\|}\right) + \frac{1}{i}(d(x, x_i)+\|y-y_i\|)$$

or equivalently

$$0 \leqq \frac{1}{k}\langle p_i, y-y_i\rangle + \|y_i-\bar{y}_i\|\,\bar{o}\left(\frac{\|y-y_i\|}{\|y_i-\bar{y}_i\|}\right)$$
$$+\frac{1}{i}\|y_i-\bar{y}_i\|^{(k-1)/k}(d(x, x_i)+\|y-y_i\|).$$

Set

$$\bar{h}_i = i^{-1/2k}\|y_i-\bar{y}_i\|^{1/k}, \qquad A_i = \frac{G(B_{\bar{h}_i}(x_i)) - y_i}{\bar{h}_i^k}$$

and observe that, by (14), $d(x_i, t_i)+\bar{h}_i \leqq \frac{1}{2i}h_i + \frac{1}{2i}h_i = \frac{1}{i}h_i \leqq h_i$. Thus $B_{\bar{h}_i}(x_i) \subset B_{h_i}(t_i)$ and

$$\forall v \in A_i, \quad 0 \leqq \frac{1}{k}\langle p_i, v\rangle + \sqrt{i}\,\bar{o}\left(\frac{\|v\|}{\sqrt{i}}\right) + i^{-(k+1)/2k}(1+\bar{h}_i^{k-1}\|v\|).$$

Hence for a sequence $\varepsilon_i \to 0+$ and for all $v \in A_i \cap MB$, $\langle p_i, v\rangle \geqq -\varepsilon_i$. Consequently

$$\forall v \in \overline{co}(A_i \cap MB), \quad \langle p_i, v\rangle \geqq -\varepsilon_i \tag{16}$$

On the other hand, by our assumptions, for all large i, $\inf_{v \in \overline{co}(A_i \cap MB)} \langle p_i, v\rangle \leqq -\rho$. But (16) yields $-\rho \geqq -\varepsilon_i$, which contradicts the choice of ε_i.

4. STABILITY OF THE UNIFORM OPEN MAPPING PRINCIPLE

The main aim of this section is to establish "stability" of the uniform open mapping principle. Namely we prove here that if a sequence of set-valued maps with closed graphs $G_i : X \to Y$ from a complete metric space X to a Banach space Y approaches uniformly a map G on a ball $B_\varepsilon(x_0)$ and satisfies the uniform open mapping principle on a neighborhood of (x_0, y_0), then so does G. This result is helpful for investigation of nonsmooth problems.

THEOREM 4.1. – *Consider a sequence of set-valued maps* $\{G_i\}_{i \geqq 0}$ *from a complete metric space* X *to a Banach space* Y *having closed graphs. Let* $y_0 \in G_0(x_0)$. *We assume that for some* $\bar{\delta}>0$ *and for every* $\lambda>0$ *there exists an integer* I_λ *such that for all* $i \geqq I_\lambda$, $x \in B_{\bar{\delta}}(x_0)$

$$G_i(x) \subset G_0(x) + \lambda B$$

If for some $0<\varepsilon<\bar{\delta}$, $\rho>0$ *and for all* $i \geqq 1$, $x \in B_\varepsilon(x_0)$, $y \in G_i(x) \cap B_\varepsilon(y_0)$ *and all* $h \in [0, \varepsilon]$

$$y + \rho h \mathring{B} \subset G_i(B_h(x))$$

then for every $(x, y) \in \text{Graph}(G_0) \cap B_{\varepsilon/4}(x_0) \times B_{\varepsilon/4}(y_0)$, $h \in \left[0, \frac{\varepsilon}{4}\right]$ *we have*

$$y + \rho h \mathring{B} \subset G_0(B_h(x)).$$

Proof. – Set $G = G_0$. It is enough to check that for every $\lambda>0$ and all (x, y), h as in the conclusion of theorem, the inclusion (9) holds true. Fix

$\lambda>0$ and assume for a moment that for some

$$(t, z) \in \operatorname{Graph}(G) \cap B_{\varepsilon/4}(x_0) \times B_{\varepsilon/4}(y_0), \qquad 0<\bar{h} \leqq \frac{\varepsilon}{4}$$

there exists $\bar{y} \in Y$ as in (10). Define $0<\Theta<1$, K as in the proof of Theorem 3.1. By Ekeland's variational principle applied to the continuous function $(x, y) \rightarrow \|y-\bar{y}\|$ on the complete metric space K with the metric

$$d_{X \times Y}((u, v), (u', v')) = d(u, u') + \frac{\lambda}{\rho}\|v-v'\| \tag{17}$$

there exists $(x, y) \in \operatorname{Graph}(G) \cap B_{\Theta\bar{h}}(t) \times B_{\Theta\bar{h}}(z)$ such that

$$\forall (u, v) \in K, \quad \|y-\bar{y}\| \leqq \|v-\bar{y}\| + \frac{\Theta\rho}{1+\lambda}\left(d(u, x) + \frac{\lambda}{\rho}\|v-y\|\right). \tag{18}$$

By the choice of $\bar{y}$, $y \neq \bar{y}$. Consider $\delta>0$ so small that

$$\Theta<\Theta+\delta<1, \eta := \frac{2\delta\rho(1+\lambda)}{2(1+\lambda)+\Theta\lambda} < \min\left\{1, \frac{\|y-\bar{y}\|}{2}, \frac{(1-\Theta)\bar{h}}{2}\right\} < \frac{\varepsilon}{4}$$

and let i be so large that for all $u \in B_{\varepsilon}(x_0)$, $G_i(u) \subset G(u) + \dfrac{\eta^2}{2(1+\lambda)} B$. Pick $v' \in G_i(x)$ such that $\|v'-y\| \leqq \dfrac{\eta^2}{2(1+\lambda)}$. Then, by (18), for all $(u, v) \in K_i := \operatorname{Graph}(G_i) \cap B_{\bar{h}}(t) \times Y$

$$\begin{aligned}\|v'-\bar{y}\| &\leqq \|v-\bar{y}\| + \frac{\Theta\rho}{1+\lambda}\left(d(u, x) + \frac{\lambda}{\rho}\|v-v'\|\right) + \frac{\eta^2}{1+\lambda} + \frac{\Theta\rho}{1+\lambda}\frac{\lambda}{\rho}\frac{\eta^2}{2(1+\lambda)} \\ &\leqq \|v-\bar{y}\| + \frac{\Theta\rho}{1+\lambda}\left(d(u, x) + \frac{\lambda}{\rho}\|v-v'\|\right) + \frac{\delta\rho\eta}{1+\lambda}\end{aligned}$$

Applying Ekeland's principle to the continuous function

$$(u, v) \rightarrow \|v-\bar{y}\| + \frac{\Theta\rho}{1+\lambda}\left(d(u, x) + \frac{\lambda}{\rho}\|v-v'\|\right)$$

on the complete metric space K_i with the metric (17), we show that for some $(\bar{u}, \bar{v}) \in K_i \cap B_\eta(x) \times B_\eta(v')$ and all $(u, v) \in K_i$

$$\begin{aligned}\|\bar{v}-\bar{y}\| + \frac{\Theta\rho}{1+\lambda}\left(d(\bar{u}, x) + \frac{\lambda}{\rho}\|\bar{v}-v'\|\right) &\leqq \|v-\bar{y}\| + \frac{\Theta\rho}{1+\lambda}(d(u, x) \\ &+ \frac{\lambda}{\rho}\|v-v'\|) + \frac{\delta\rho}{1+\lambda}\left(d(\bar{u}, u) + \frac{\lambda}{\rho}\|\bar{v}-v\|\right) \leqq \|v-\bar{y}\|\end{aligned}$$

$$+\frac{\Theta\rho}{1+\lambda}\left(d(\bar{u},x)+\frac{\lambda}{\rho}\|\bar{v}-v'\|\right)$$
$$+\frac{\Theta\rho}{1+\lambda}\left(d(\bar{u},u)+\frac{\lambda}{\rho}\|\bar{v}-v\|\right)+\frac{\delta\rho}{1+\lambda}\left(d(\bar{u},u)+\frac{\lambda}{\rho}\|\bar{v}-v\|\right).$$

Hence for every $(u, v)\in K_i$

$$\|\bar{v}-\bar{y}\|\leqq\|v-\bar{y}\|+(\Theta+\delta)\frac{\rho}{1+\lambda}\left(d(\bar{u},u)+\frac{\lambda}{\rho}\|\bar{v}-v\|\right).$$

Moreover from the choice of η, v'

$$\|\bar{v}-\bar{y}\|\geqq\|y-\bar{y}\|-\|v'-y\|-\|v'-\bar{v}\|\geqq\frac{\|y-\bar{y}\|}{4}>0.$$

Observe that $d(\bar{u}, x_0)\leqq\frac{3}{4}\varepsilon$, $\|\bar{v}-y_0\|\leqq\varepsilon$ and, by the assumptions of theorem, for all $h\in[0,\varepsilon]$

$$\bar{v}+\rho h\frac{\bar{y}-\bar{v}}{\|\bar{y}-\bar{v}\|}\in G_i(B_h(\bar{u})).$$

Since $B_\eta(x)\subset\mathring{B}_{\bar{h}}(t)$, for all small $h>0$

$$\|\bar{v}-\bar{y}\|\leqq\left\|\bar{v}+\rho h\frac{\bar{y}-\bar{v}}{\|\bar{y}-\bar{v}\|}-\bar{y}\right\|+(\Theta+\delta)\frac{\rho}{1+\lambda}\left(h+\frac{\lambda}{\rho}h\rho\right)$$
$$\leqq\|\bar{v}-\bar{y}\|\left(1-\rho h\frac{1}{\|\bar{y}-\bar{v}\|}\right)+(\Theta+\delta)\frac{\rho h}{1+\lambda}(1+\lambda)$$
$$\leqq\|\bar{v}-\bar{y}\|-\rho h+(\Theta+\delta)\rho h.$$

Thus $\rho\leqq(\Theta+\delta)\rho<\rho$. The obtained contradiction ends the proof. □

Altough it is possible to prove stability of the higher order uniform open mapping principle in uniformly smooth Banach spaces using ideas from Section 3 and from the above proof we do not do it here: our applications use only the above first order result.

5. VARIATIONS OF SET-VALUED MAPS

To check whether the uniform open mapping principle holds true it is convenient to introduce variations of set-valued maps which measure "infinitesimal changes" of the map. We recall first the notions of Kuratowski's lim sup and lim inf [3]:

Let T be a metric space $A_\tau \subset Y$, $\tau \in T$ be a family of subsets of Y. The Kuratowski lim sup and lim inf of A_τ at τ_0 are closed sets given by

$$\limsup_{\tau \to \tau_0} A_\tau = \{ v \in Y \mid \liminf_{\tau \to \tau_0} \operatorname{dist}(v, A_\tau) = 0 \}$$
$$\liminf_{\tau \to \tau_0} A_\tau = \{ v \in Y \mid \lim_{\tau \to \tau_0} \operatorname{dist}(v, A_\tau) = 0 \}.$$

Consider a metric space X and a Banach space Y. Let $G: X \to Y$ be a set-valued map, that is for all $x \in X$, $G(x)$ is a (possibly empty) subset of Y.

DEFINITION 5.1. — *Let* $(x, y) \in \operatorname{Graph}(G)$, $k>0$.

(i) *The contingent variation of* G *at* (x, y) *is the closed subset of* Y

$$G^{(1)}(x, y) := \limsup_{h \to 0+} \frac{G(B_h(x)) - y}{h}$$

(ii) *The k-th order variation of* G *at* (x, y) *is the closed subset of* Y

$$G^k(x, y) := \liminf_{\substack{(x', y') \to_G (x, y) \\ h \to 0+}} \frac{G(B_h(x')) - y'}{h^k}$$

where $\to_G$ *denotes the convergence in* Graph (G).

In other words $v \in G^{(1)}(x, y)$ if and only if there exist sequences $h_i \to 0+$, $v_i \to v$ such that $y + h_i v_i \in G(B_{h_i}(x))$. The word contingent is used because the definition reminds that of the contingent cone of Bouligand.

Similarly $v \in G^k(x, y)$ if and only if for all sequences $h_i \to 0+$, $(x_i, y_i) \to_G (x, y)$ there exists a sequence $v_i \to v$ such that $y_i + h_i^k v_i \in G(B_{h_i}(x_i))$.

Clearly, $G^{(1)}(x, y)$ and $G^k(x, y)$ are closed sets starshaped at zero. When X is a Banach space, $G: X \to Y$ is a Gâteaux differentiable at some $x \in X$ function and B denotes the closed unit ball in X, then $\overline{G'(x)(B)} \subset G^{(1)}(x, G(x))$. If G is Fréchet differentiable at x then $\overline{G'(x)(B)} = G^{(1)}(x, G(x))$. Moreover if G is continuously differentiable at x then $\overline{G'(x)(B)} = G^1(x, G(x))$.

Remark. — Observe that if G verifies the uniform open mapping principle at (x_0, y_0) in the sense that for some $\rho>0$, $k>0$ and for all $(x, y) \in \operatorname{Graph}(G)$ near (x_0, y_0) and for all small $h>0$

$$y + \rho h^k \mathring{B} \subset G(B_h(x))$$

then $\rho B \subset G^k(x_0, y_0)$. Moreover if $k=1$ then for some $\varepsilon>0$

$$\rho B \subset \bigcap_{\substack{(x, y) \in \operatorname{Graph}(G) \\ (x, y) \in B_\varepsilon(x_0) \times B_\varepsilon(y_0)}} G^{(1)}(x, y)$$

In Sections 6 and 7 we show that the converse statement holds true for the first order condition. For the high order one we have to impose some additional assumptions. □

Remark. — When X is a Banach space the following two set-valued derivatives $DG(x, y)$, $CG(x, y)$ of G at a point $(x, y) \in \text{Graph}(G)$ were considered in [3], Chapter 5:

$$\forall u \in X, \quad v \in DG(x, y)u \Leftrightarrow (u, v) \in T_{\text{Graph}(G)}(x, y)$$
$$\forall u \in X, \quad v \in CG(x, y)u \Leftrightarrow (u, v) \in C_{\text{Graph}(G)}(x, y)$$

where $T_{\text{Graph}(G)}(x, y)$ and $C_{\text{Graph}(G)}(x, y)$ denote respectively the contingent and Clarke's tangent cones to Graph (G) at (x, y) (*see* [9], [3]). It is not difficult to show that

$$DG(x, y)(B) \subset G^{(1)}(x, y), \qquad CG(x, y)(B) \subset G^1(x, y).$$

Moreover if for a subset $K \subset Y$, $G(.) \equiv K$, then for every $x \in X$, $y \in K$ we have $G^{(1)}(x, y) = T_K(y)$, $G^1(x, y) = C_K(y)$. In this way some results from [3], Chapter 5 and [4], which use the set-valued derivatives $CG(x, y)$, $DG(x, y)$ are consequences of those proved in Section 6.

Let $co\,(\overline{co})$ denote the convex (closed convex) hull. We proved in [18] some properties of high order variations. In particular, if $Y = \mathbf{R}^n$ then for any integer $k \geqq 1$

$$(n+1)^{1-k}\,\overline{co}\,G^k(x, y) \subset G^k(x, y).$$

The following theorem extends those results to the infinite dimensional case.

THEOREM 5.2. — *For every* $(x, y) \in \text{Graph}(G)$, $k > 0$ *we have:*

(i) *For all* $K \geqq k$, $0 \in G^k(x, y) \subset G^K(x, y)$;

(ii) *For all* $s > 0$, $\mathbf{R}_+ G^k(x, y) \subset G^{k+s}(x, y)$;

(iii) *For all* $\lambda_i \geqq 0$, $v_i \in G^k(x, y)$, $i = 0, \ldots, m$ *satisfying* $\sum_{i=0}^{m} \lambda_i = 1$,

$$\sum_{i=0}^{m} \lambda_i^k v_i \in G^k(x, y);$$

(iv) *For all* $v \in co\,G^k(x, y)$ *there exists* $\varepsilon > 0$ *such that* $\varepsilon v \in G^k(x, y)$;

(v) $\bigcup_{\lambda \geqq 0} \lambda\, co\, G^k(x, y) = \bigcup_{\lambda \geqq 0} \lambda G^k(x, y)$;

(vi) $\bigcup_{\lambda \geqq 0} \lambda G^k(x, y) = Y \Leftrightarrow 0 \in \text{Int}\, co\, G^k(x, y)$. *Moreover if* $Y = \mathbf{R}^n$ *these conditions are equivalent to*

$$\exists v_1, \ldots, v_p \in G^k(x, y) \qquad \textit{such that} \quad 0 \in \text{Int}\, co\{v_1, \ldots, v_p\}$$

Proof. — For all $(x', y') \in \text{Graph}(G)$ and $h > 0$ we have $y' \in G(B_h(x'))$. Therefore $0 \in G^k(x, y)$. Fix $K > k$. Then for all $h \in]0, 1]$, $h^{K/k} \leqq h$. Hence $G(B_{h^{K/k}})(x')) - y' \subset G(B_h(x')) - y'$.

Consequently for all $v \in Y$

$$\operatorname{dist}\left(v, \frac{G(B_h(x')) - y'}{h^K}\right) \leqq \operatorname{dist}\left(v, \frac{G(B_{h(K/k)}(x')) - y'}{(h^{K/k})^k}\right)$$

and (i) follows. To prove (ii) fix $s>0$, $v \in G^k(x, y)$, $\lambda \geqq 0$ and set $\bar{k}=k+s$. Let $(x_i, y_i) \to_G (x, y)$, $h_i \to 0+$, $h'_i = \lambda^{1/k} h_i^{\bar{k}/k}$. Then for all large i, $h'_i \leqq h_i$. Let $v_i \to v$ be such that $y_i + h'^k_i v_i = y_i + h_i^{\bar{k}} \lambda v_i \in G(B_{h'_i}(x_i)) \subset G(B_{h_i}(x_i))$. This implies that $\lambda v \in G^{\bar{k}}(x, y)$ and since $\lambda \geqq 0$ is arbitrary (ii) follows. Fix next λ_i, v_i as in (iii). We proceed by induction. Observe first that

$$\lim_{\substack{(u,\, z) \to_G (x,\, y) \\ h \to 0+}} \operatorname{dist}\left(\lambda_0^k v_0, \frac{G(B_{\lambda_0 h}(u)) - z}{h^k}\right) = 0$$

Assume that we already proved that for some $0 \leqq s < m$

$$\lim_{\substack{(u,\, z) \to_G (x,\, y) \\ h \to 0+}} \operatorname{dist}\left(\sum_{i=0}^{s} \lambda_i^k v_i, \frac{G(B_{h(\lambda_0 + \ldots + \lambda_s)}(u)) - z}{h^k}\right) = 0. \tag{19}$$

Fix $(u_j, z_j) \to_G (x, y)$, $h_j \to 0+$. By (19), for some $x_j \in B_{h_j(\lambda_0 + \ldots + \lambda_s)}(u_j)$ and $w_j \to \sum_{i=0}^{s} \lambda_i^k v_i$, we have $z_j + h_j^k w_j \in G(x_j)$. This and Definition 5.1 imply that

$$\lim_{j \to \infty} \operatorname{dist}\left(\lambda_{s+1}^k v_{j+1}, \frac{G(B_{h_j \lambda_{s+1}}(x_j)) - z_j - h_j^k w_j}{h_j^k}\right) = 0 \tag{20}$$

Since $B_{h_j \lambda_{s+1}}(x_j) \subset B_{h_j(\lambda_0 + \ldots + \lambda_{s+1})}(u_j)$, the definition of w_j and (20) yield

$$\lim_{j \to \infty} \operatorname{dist}\left(\sum_{i=0}^{s+1} \lambda_i^k v_i, \frac{G(B_{h_j(\lambda_0 + \ldots + \lambda_{s+1})}(u_j)) - z_j}{h_j^k}\right) = 0$$

Because the sequences $(u_j, z_j) \to_G (x, y)$, $h_j \to 0+$ are arbitrary we proved that (19) is verified with s replaced by $s+1$. Applying (19) with $s=m$ and using the identity $\sum_{i=0}^{m} \lambda_i = 1$ we obtain (iii). Fix next $v \in co\, G^k(x, y)$ and let $\mu_i > 0$, $v_i \in G^k(x, y)$ be such that $\sum_{i=0}^{m} \mu_i = 1$, $\sum_{i=0}^{m} \mu_i v_i = v$. Set $\lambda_i = (\mu_i)^{1/k}/(m+1)$ for $i=0, \ldots, m$ and $\lambda_{m+1} = 1 - \sum_{i=0}^{m} \lambda_i$, $v_{m+1} = 0$. Since $\mu_i \leqq 1$ we have $\sum_{i=0}^{m} \lambda_i \leqq 1$ and therefore, by (iii) and (i), $\sum_{i=0}^{m+1} \lambda_i^k v_i \in G^k(x_0, y_0)$. On the other hand $(m+1)^{-k} v = \sum_{i=0}^{m} ((m+1)^{-1} \mu_i^{1/k})^k v_i = \sum_{i=0}^{m+1} \lambda_i^k v_i$ and therefore (iv) holds true with $\varepsilon = (m+1)^{-k}$. The statement (v) follows from (iv). Assume next that $\bigcup_{\lambda \geqq 0} \lambda G^k(x, y) = Y$. Since $G^k(x, y)$ is starshaped at zero, for all $0 \leqq \lambda \leqq \mu$, $\lambda G^k(x, y) \subset \mu G^k(x, y)$. Thus $\bigcup_{\lambda \in Z} \lambda G^k(x, y) = Y$, where Z denote the set of positive integers. Using Baire's theorem, we prove that $G^k(x, y)$ has a nonempty interior and therefore also $co\, G^k(x, y)$ does. Assume for a moment that zero is not an interior point of $co\, G^k(x, y)$. By the separation theorem, there exists a non zero

$$p \in (co\, G^k(x, y))^+ = (\bigcup_{\lambda \geqq 0} \lambda\, co\, G^k(x, y))^+ \subset (\bigcup_{\lambda \geqq 0} \lambda G^k(x, y))^+ = \{0\}.$$

The obtained contradiction proves that $0 \in \operatorname{Int} co\, G^k(x, y)$. The converse follows from (v). The last statement is a consequence of the Caratheodory theorem. □

Remark. — Observe that the proof of (iv) and Caratheodory's theorem yield that whenever $\dim Y < \infty$

$$(\dim Y + 1)^{-k}\, \overline{co}\, G^k(x, y) \subset G^k(x, y).$$

6. FIRST ORDER INVERSE MAPPING THEOREMS

As one may expect, the first order results are more simple and require less assumptions that their high order analogous. This is why we study them separately.

In this section, we assume that X is a complete metric space with the metric d and Y is a Banach space with the norm $\| \cdot \|$. Consider a set-valued map $G : X \to Y$ having a closed graph.

THEOREM 6.1. — *Let $y_0 \in G(x_0)$ and assume that for some $\varepsilon > 0$, $\rho > 0$*

$$\rho B \subset \bigcap_{\substack{(x,\, y) \in \operatorname{Graph}(G) \\ d(x,\, x_0) \leqq \varepsilon,\ \| y - y_0 \| \leqq \varepsilon}} G^{(1)}(x, y). \tag{21}$$

Then for every $(x_1, y_1) \in \operatorname{Graph}(G) \cap B_{\varepsilon/4}(x_0) \times B_{\varepsilon/4}(y_0)$, $y_2 \in Y$ satisfying $\| y_2 - y_1 \| < \min\left\{\frac{\varepsilon}{8}, \frac{\varepsilon\rho}{4}\right\}$

$$\operatorname{dist}(x_1, G^{-1}(y_2)) \leqq \frac{1}{\rho} \| y_1 - y_2 \|. \tag{22}$$

Remark. – Inequality (22) implies that for every $L>\frac{1}{\rho}$, G^{-1} is pseudo-Lipschitz at (y_0, x_0) with the Lipschitz constant L (*see* [1]). □

COROLLARY 6.2. – *Consider* $y_0 \in G(x_0)$, $\rho>0$. *Then* (22) *holds true for all* $(x_1, y_1, y_2) \in \mathrm{Graph}(G)\times Y$ *near* (x_0, y_0, y_0) *if and only if for some* $\varepsilon>0$ *inclusion* (21) *is verified.*

Proof. – Indeed if for all $(x_1, y_1, y_2) \in \mathrm{Graph}(G)\times Y$ near (x_0, y_0, y_0) we have (22), then for every such (x_1, y_1, y_2) and every $0<\rho'<\rho$

$$\mathrm{dist}(x_1, G^{-1}(y_2)) < \frac{1}{\rho'}\|y_1-y_2\|$$

and therefore for some $x' \in G^{-1}(y_2)$, $d(x_1, x') \leqq \frac{1}{\rho'}\|y_1-y_2\|$. This implies that for all small $h>0$, $y_1+\rho' h B \subset G(B_h(x_1))$. Hence $\rho' B \subset G^{(1)}(x_1, y_1)$. Thus (21) is satisfied for some $\varepsilon>0$ and every $\rho'<\rho$. Since the right-hand side of (21) is closed we deduce that the inclusion (21) is verified as well with $\rho=\rho'$. The opposite follows from Theorem 6.1. □

The above contains a classical result of functional analysis:

COROLLARY 6.3. – *Let* $g: X \to Y$ *be a function between two Banach spaces. Assume that g is continuously differentiable at some* $x_0 \in X$ *and for some* $\rho>0$

$$\rho B \subset \overline{g'(x_0) B} \tag{23}$$

Then for all (x, y) *near* (x_0, y_0), $\mathrm{dist}(x, g^{-1}(y)) \leqq \frac{2}{\rho}\|g(x)-y\|$. *In particular for* $y_0=g(x_0)$ *and for all* $x \in x_0+\mathrm{Ker}\, g'(x_0)$

$$\mathrm{dist}(x, g^{-1}(y_0)) = o(\|x-x_0\|)$$

and thus the tangent space to $g^{-1}(y_0)$ *at* x_0 *coincides with* $\mathrm{Ker}\, g'(x_0)$.

To derive such result, it is enough to observe that since $g \in C^1$ locally at x_0 by our assumption, for all x near x_0, $\frac{\rho}{2} B \subset \overline{g'(x) B} \subset g^{(1)}(x, g(x))$.

Remark. – By the Banach open mapping theorem, assumption (23) is verified whenever $g'(x_0)$ is surjective, *i. e.* $g'(x_0) X = Y$, □

Proof of Theorem 6.1. – By Theorem 2.2, it is enough to check that for every $\lambda>0$ and for all

$$(x, y) \in \mathrm{Graph}(G) \cap B_{\varepsilon/2}(x_0) \times B_{\varepsilon/2}(y_0),\ h \in \left[0, \frac{\varepsilon}{2}\right],$$

inclusion (9) holds true. Fix $\lambda>0$, $(t, z)\in \operatorname{Graph}(G)\cap B_{\varepsilon/2}(x_0)\times B_{\varepsilon/2}(y_0)$, $0<\bar{h}\leqq\frac{\varepsilon}{2}$ and assume for a moment that there exists $\bar{y}\in Y$ satisfying (10). Define $0<\Theta<1$ and K as in the proof of Theorem 3.1. Applying the Ekeland variational principle [13] to the continuous function $(x, y)\to \|y-\bar{y}\|$ on the complete metric space K with the metric given by (17), we prove the existence of $(x, y)\in B_{\Theta\bar{h}}(t)\times B_{\Theta\bar{h}}(z)$ verifying (18). Observe that $\bar{x}\in \operatorname{Int} B_{\bar{h}}(t)$, $\bar{y}\in \operatorname{Int} B_{\varepsilon/2}(z)$ and $y\neq\bar{y}$. Set $w=-\rho(y-\bar{y})/\|y-\bar{y}\|$. By our assumption there exist $h_i\to 0+$, $w_i\to w$ such that $y+h_i w_i\in G(B_{h_i}(x))$. Hence, from (18) we deduce that for all large i

$$\|y-\bar{y}\|\leqq\|y+h_i w-\bar{y}\|+h_i\|w_i-w\|+\frac{\Theta\rho}{1+\lambda}h_i\left(1+\frac{\lambda}{\rho}\|w_i\|\right)$$
$$=\left(1-h_i\frac{\rho}{\|y-\bar{y}\|}\right)\|y-\bar{y}\|+h_i\|w_i-w\|+\frac{\Theta\rho}{1+\lambda}h_i\left(1+\frac{\lambda}{\rho}\|w_i\|\right)$$

and therefore $h_i\rho\leqq h_i\|w_i-w\|+\frac{\Theta\rho}{1+\lambda}h_i\left(1+\frac{\lambda}{\rho}\|w_i\|\right)$. Dividing by ρh_i and taking the limit yields $1\leqq\Theta$. The obtained contradiction ends the proof. □

THEOREM 6.4 (A characterization of the image). – *Let* $y_0\in G(x_0)$. *Assume that there exist closed convex subsets* $K(x, y)\subset G^{(1)}(x, y)$, $\varepsilon>0$ *and a compact set* $Q\subset Y$ *such that*

$$\operatorname{Int}\bigcap_{\substack{(x,\, y)\in \operatorname{Graph}(G)\\ (x,\, y)\in B_\varepsilon(x_0)\times B_\varepsilon(y_0)}}(K(x, y)+Q)\neq\varnothing. \tag{24}$$

Then at least one of the following two statements holds true:

(i) *There exist* $L>0$, $\delta>0$ *such that for all*

$$(x_1, y_1, y_2)\in(\operatorname{Graph}(G)\cap B_\delta(x_0)\times B_\delta(y_0))\times B_\delta(y_0)$$
$$\operatorname{dist}(x_1, G^{-1}(y_2))\leqq L\|y_1-y_2\|.$$

(ii) *There exists a non zero* $p\in Y^*$ *such that*

$$\forall w\in \liminf_{(x,\, y)\to_G (x_0,\, y_0)} K(x, y),\qquad \langle p, w\rangle\geqq 0. \tag{25}$$

Consequently if for some $\delta>0$, y_0 *is a boundary point of* $G(B_\delta(x_0))$, *then there exists a non zero* $p\in Y^*$ *such that* (25) *is satisfied.*

Observe that if Y is a finite dimensional space and $0\in K(x, y)$ for all $(x, y)\in \operatorname{Graph}(G)$, then condition (24) is always verified with Q equal to the closed unit ball.

Proof. – If for some $\varepsilon'>0$

$$0 \in \operatorname{Int} \bigcap_{\substack{(x,\, y) \in \operatorname{Graph}(G) \\ (x,\, y) \in B_{\varepsilon'}(x_0) \times B_{\varepsilon'}(y_0)}} K(x, y)$$

then Theorem 6.1 implies statement (i). Otherwise if for every $\varepsilon'>0$ the above is not satisfied, by the separation theorem for a sequence $(x_i, y_i) \to_G (x_0, y_0)$ and some $p_i \in Y^*$ of $\|p_i\|_{Y^*}=1$, $\mu_i \to 0+$

$$\inf_{k \in K(x_i,\, y_i)} \langle p_i, k \rangle \geqq -\mu_i. \tag{26}$$

Let $z \in Y$, $\rho>0$ be such that the ball $z+\rho B$ is contained in the left hand side of (24). Pick $w_i \in B$ such that $\langle p_i, w_i \rangle \geqq 1-\frac{1}{i}$ and let $k_i \in K(x_i, y_i)$, $q_i \in Q$ be such that $z-\rho w_i = k_i + q_i$.

Then

$$\langle p_i, z-q_i \rangle = \langle p_i, \rho w_i \rangle + \langle p_i, k_i \rangle \geqq \rho\left(1-\frac{1}{i}\right)-\mu_i$$

Consider a subsequence q_{i_j} converging to some $q \in Q$. Then for all large j

$$\langle p_{i_j}, z-q \rangle \geqq \frac{\rho}{2}$$

Let $p \in Y^*$ be a weak-$*$ cluster point of $\{p_{i_j}\}$. Passing to the limit in the last inequality we get $\langle p, z-q \rangle \geqq \rho/2$. Therefore p is different from zero. On the other hand the inequality (26) yields (25). □

When the norm of Y is differentiable, then a stronger result may be proved:

THEOREM 6.5. – *Assume that the norm of* Y *is Gâteaux differentiable away from zero and let* $y_0 \in G(x_0)$. *If for some* $\varepsilon>0$, $\rho>0$, $M>0$

$$\rho B \subset \bigcap_{\substack{(x,\, y) \in \operatorname{Graph}(G) \\ (x,\, y) \in B_{\varepsilon}(x_0) \times B_{\varepsilon}(y_0)}} \overline{co}(G^{(1)}(x, y) \cap MB) \tag{27}$$

then for every $(x_1, y_1) \in \operatorname{Graph}(G) \cap B_{\varepsilon/4}(x_0) \times B_{\varepsilon/4}(y_0)$, $y_2 \in Y$ *satisfying* $\|y_2 - y_1\| \leqq \min\left\{\frac{\varepsilon}{8}, \frac{\varepsilon\rho}{4}\right\}$

$$\operatorname{dist}(x_1, G^{-1}(y_2)) \leqq \frac{1}{\rho}\|y_1 - y_2\|.$$

Remark. – It was shown in [17] that in the case when G is single-valued, the constant M in the assumption (27) may be taken equal to $+\infty$. □

COROLLARY 6.6. — *Assume that* $Y=\mathbf{R}^n$ *and that for some* $M>0$

$$0\in \operatorname{Int} \liminf_{(x,\, y)\to_G (x_0,\, y_0)} \overline{co}\,(G^{(1)}(x, y)\cap MB).$$

Then there exist $\varepsilon>0$, $\rho>0$ *such that conclusions of Theorem* 6.5 *are valid.*

Proof. — By Theorem 2.2 it is enough to show that (9) holds true for all $\lambda>0$, $(x, y)\in \operatorname{Graph}(G)\cap B_{\varepsilon/2}(x_0)\times B_{\varepsilon/2}(y_0)$, $0<\bar{h}\leqq\frac{\varepsilon}{2}$. Fix such λ, $\bar{h}$, $(t, z)=(x, y)$ and assume for a while that there exists $\bar{y}\in Y$ satisfying (10). Define Θ, K as in the proof of Theorem 3.1. Applying the Ekeland variational principle [13] to the continuous function $(x, y)\to\|y-\bar{y}\|$ on the complete metric space K with the metric (11), we prove that for some $(x, y)\in B_{\Theta\bar{h}}(t)\times B_{\Theta\bar{h}}(z)$, inequality (12) is verified. Since $y\neq\bar{y}$, from differentiability of the norm, we infer the existence of $p\in Y^*$ of $\|p\|_{Y^*}=1$ such that for all $h_j\to 0+$, $v_j\to v$

$$\|y+h_j v_j-\bar{y}\|=\|y-\bar{y}\|+\langle p, h_j v\rangle+o_v(h_j) \tag{28}$$

where $\liminf_{j\to\infty} o_v(h_j)/h_j=0$. Fix $v\in G^{(1)}(x,y)$ and let $h_j\to 0+$, $v_j\to v$ be such that $y+h_j v_j\in G(B_{h_j}(x))$. Then from (12) and (28)

$$0\leqq\langle p, h_j v\rangle+\frac{\Theta\rho}{1+\lambda}\left(h_j+\frac{\lambda}{M}h_j\|v_j\|\right)+o_v(h_j)$$

Dividing by h_j and taking the limit yields

$$\langle p, v\rangle\geqq-\frac{\Theta\rho}{1+\lambda}\left(1+\frac{\lambda}{M}\|v\|\right)$$

and therefore

$$\forall v\in\overline{co}\,(G^{(1)}(\bar{x}, \bar{y})\cap MB),\quad \langle p, v\rangle\geqq-\Theta\rho. \tag{29}$$

Since $(x, y)\in B_\varepsilon(x_0)\times B_\varepsilon(y_0)$, by the assumption of theorem, $\rho B\subset\overline{co}\,(G^{(1)}(x, y)\cap MB)$. Using that $0<\Theta<1$ we derive a contradiction. □

The following theorem provides a stronger sufficient condition for local invertibility but does not allow to estimate the Lipschitz constant.

THEOREM 6.7. — *Assume that the norm of* Y *is Gâteaux differentiable away from zero. Let* $y_0\in G(x_0)$. *Further assume that there exist* $\varepsilon>0$, $M>0$ *and a compact* $Q\subset Y$ *such that*

$$\operatorname{Int}\bigcap_{\substack{(x,\, y)\in \operatorname{Graph}(G)\\ (x,\, y)\in B_\varepsilon(x_0)\times B_\varepsilon(y_0)}}(\overline{co}\,(G^{(1)}(x, y)\cap MB)+Q)\neq\varnothing. \tag{30}$$

Then the following statements are equivalent

(i) $\left(\liminf_{(x,\,y)\to_G (x_0,\,y_0)} \overline{co}(G^{(1)}(x,y)\cap MB)+G^1(x_0,y_0)\right)^+=\{0\}$

(ii) *for some* $\delta>0$, $L>0$ *and for all*

$$(x_1, y_1, y_2)\in(\text{Graph}(G)\cap B_\delta(x_0)\times B_\delta(y_0))\times B_\delta(y_0)$$
$$\text{dist}(x_1, G^{-1}(y_2))\leqq L\|y_1-y_2\|$$

In particular if for some $\delta>0$, y_0 *is a boundary point of* $G(B_\delta(x_0))$, *then there exists a non zero* $p\in Y^*$ *such that*

$$\forall w\in \liminf_{(x,\,y)\to_G (x_0,\,y_0)} \overline{co}(G^{(1)}(x,y)\cap MB)+G^1(x_0,y_0),\quad \langle p, w\rangle\geqq 0 \quad (31)$$

Moreover if G *is single valued, then* M *in* (30), (31) *and* (i) *may be taken equal to* $+\infty$.

As a consequence we deduce a very useful criterion using interiority properties in subspaces with finite codimension.

COROLLARY 6.8. — *Let* $y_0\in G(x_0)$. *Assume that* Y *is a Hilbert space and that there exist a closed subspace* $H\subset Y$ *of finite codimension,* $\rho>0$, $M>0$, $z\in Y$ *such that for all* $(x,y)\in\text{Graph}(G)$ *near* (x_0,y_0)

$$z+\rho B_H\subset\overline{co}(G^{(1)}(x,y)\cap MB)$$

where B_H *denote the closed unit ball in* H. *Then the conclusions of Theorem* 6.7 *hold true.*

Proof. — Let P denote the orthogonal projection onto H and Q′ denote the closed unit ball in the ortogonal space $H^\perp$. Then the set $Q:=\rho Q'$ is compact and for all $x\in\rho B$, $x=x-Px+Px\in H^\perp+H$, where $\|Px\|\leqq\rho$ and $\|x-Px\|\leqq\rho$. Hence $\rho B\subset\rho B_H+\rho Q'$ and

$$z+\rho B\subset z+\rho B_H+\rho Q'\subset\overline{co}(G^{(1)}(x,y)\cap MB)+\rho Q'$$

for all $(x,y)\in\text{Graph}(G)$ near (x_0,y_0). Therefore (30) is verified and the result follows from Theorem 6.7. □

Proof. — By Theorem 2.1, (ii) implies that for all $(x,y)\in\text{Graph}(G)$ near (x_0,y_0) and all small $h>0$

$$y+\frac{1}{L}h\mathring{B}\subset G(B_h(x))$$

and from Definition 5.1 we deduce that $\frac{1}{L}B\subset G^1(x_0,y_0)$ which implies (i). To show that (i) ⇒ (ii), by Theorem 2.1, it is enough to prove that for some $\rho>0$ and all $(x,y)\in\text{Graph}(G)$ near (x_0,y_0) and all small $h>0$

$$y+\rho h\mathring{B}\subset G(B_h(x))$$

Assume for a moment that for some $(t_i, z_i) \to_G (x_0, y_0)$, $\bar{h}_i \to 0+$ there exist

$$\bar{y}_i \notin G(B_{\bar{h}_i}(t_i)), \qquad \|\bar{y}_i - z_i\| < \frac{\bar{h}_i}{i^2}, \qquad i=1,2,\ \ldots \tag{32}$$

We shall derive a contradiction. Set $K_i = \text{Graph}(G) \cap B_{\bar{h}_i}(t_i) \times Y$. We apply the Ekeland variational principle ([12], [13]) to the continuous functions $K_i \ni (x, y) \to \|y - \bar{y}_i\|$, $i=1,2, \ldots$ to prove the existence of $(x_i, y_i) \in \text{Graph}(G) \times B_{\bar{h}_{i/i}}(t_i) \times B_{\bar{h}_{i/i}}(z_i)$ such that for all $(x, y) \in K_i$

$$\|y_i - \bar{y}_i\| \leqq \|y - \bar{y}_i\| + \frac{1}{i}(d(x, x_i) + \|y - y_i\|). \tag{33}$$

By differentiability of the norm and by (32), there exist $p_i \in Y^*$, $\|p_i\|_{Y^*} = 1$ such that for all $h_j \to 0+$, $v_j \to v$, we have

$$\|y_i + h_j v_j - \bar{y}_i\| = \|y_i - \bar{y}_i\| + \langle p_i, h_j v \rangle + o_{i,\,v}(h_j)$$

where $\lim_{h \to 0+} o_{i,\,v}(h)/h = 0$. Fix $v \in G^{(1)}(x_i, y_i)$ and let $h_j \to 0+$, $v_j \to v$ be such that $y_i + h_j v_j \in G(B_{h_j}(x_i))$. Setting $y = y_i + h_j v_j$ in (33) we obtain

$$0 \leqq \langle p_i, h_j v \rangle + o_{i,\,v}(h_j) + \frac{1}{i}(h_j + h_j \|v_j\|). \tag{34}$$

Dividing by h_j and taking the limit when $j \to \infty$ yields that for all $v \in G^{(1)}(x_i, y_i)$, $\langle p_i, v \rangle \geqq -\frac{1}{i}(1 + \|v\|)$ and therefore

$$\forall v \in \overline{co}(G^{(1)}(x_i, y_i) \cap MB), \quad \langle p_i, v \rangle \geqq -\frac{1}{i}(1 + M) \tag{35}$$

By (30) there exist $z \in Y$, $\rho > 0$ and $a_i \in co(G^{(1)}(x_i, y_i) \cap MB)$, $q_i \in Q$, $w_i \in B$ such that $\langle p_i, w_i \rangle \geqq 1 - \frac{1}{i}$, $z - \rho w_i = a_i + q_i$. From (35), $\langle p_i, z - \rho w_i - q_i \rangle \geqq -\frac{1+M}{i}$. Consider a subsequence q_{i_j} converging to some $q \in Q$. Then for all large j

$$\langle p_{i_j}, z - q \rangle \geqq \frac{\rho}{2}.$$

Let $p \in Y^*$ be a weak-$*$ cluster point of $\{p_{i_j}\}$. Then from the last inequality we deduce that $p \neq 0$. On the other hand (35) implies that

$$\forall w \in \liminf_{(x,\, y) \to_G (x_0,\, y_0)} \overline{co}(G^{(1)}(x, y) \cap MB), \quad \langle p, w \rangle \geqq 0 \tag{36}$$

Fix next $v\in G^1(x_0, y_0)$ and choose $h_i \to 0+$ in such way that

$$\lim_{i\to\infty} \frac{o_{i,\,v}(h_i)}{h_i}=0 \qquad \text{and} \qquad B_{h_i}(x_i)\subset B_{\bar{h_i}}(t_i).$$

Let $v_i \to v$ be such that $y_i+h_i v_i\in G(B_{h_i}(x_i))$. Then from (34) there exist $\varepsilon_i \to 0+$ such that

$$\langle p_i, v\rangle \geqq -\varepsilon_i-\frac{1}{i}(1+\|v_i\|).$$

Since p is a weak-$*$ cluster point of $\{p_i\}$, we infer from the last inequality that for all $v\in G^1(x_0, y_0)$, $\langle p, v\rangle \geqq 0$. This, (36) and (i) together yield that $p=0$. The obtained contradiction proves (ii). When G is single-valued, the Ekeland principle has to be applied in the same way as before to the continuous functions $x\to\|G(x)-\bar{y}_i\|$ on the complete metric spaces $B_{\bar{h_i}}(t_i)$.

7. HIGH ORDER INVERSE MAPPING THEOREMS

We impose here on G, X the same assumptions than in Section 6. To prove a high order inverse function theorem we need more smoothness on the space Y. Namely we assume in this section that Y is uniformly smooth (*see* Section 3). Observe that Theorems 3.2 and 2.1 immediatly yield.

THEOREM 7.1. – *Let* $y_0\in G(x_0)$ *and assume that for some* $k\geqq 1$, $\rho>0$, $M>0$ *and all* $(x, y)\in \mathrm{Graph}(G)$ *near* (x_0, y_0) *and all small* $h>0$

$$\rho B\subset \overline{co}\left(\frac{G(B_h(x))-y}{h^k}\cap MB\right)$$

or equivalently

$$\inf_{\|p\|_{Y^*}=1}\left\{\langle p, v\rangle \,|\, v\in\frac{G(B_h(x))-y}{h^k}\cap MB\right\}\geqq \rho.$$

Then there exists $L>0$ *such that for all* $(x_1, y_1)\in \mathrm{Graph}(G)$ *near* (x_0, y_0) *and all* $y_2\in Y$ *near* y_0

$$\mathrm{dist}(x_1, G^{-1}(y_2))\leqq L\|y_1-y_2\|^{1/k}.$$

THEOREM 7.2 (High Order Inverse Function Theorem). – *Let* $y_0\in G(x_0)$ *and assume that* (30) *holds true for some* $\varepsilon>0$, $M\geqq 0$ *and a compact set* $Q\subset Y$. *If for some* $k\geqq 1$, $0\in \mathrm{Int}\,\overline{co}\,G^k(x_0, y_0)$, *then there exists* $L>0$ *such that for all* $(x_1, y_1)\in \mathrm{Graph}(G)$ *near* (x_0, y_0) *and for all* $y_2\in Y$

near y_0

$$\operatorname{dist}(x_1, G^{-1}(y_2)) \leqq L \|y_1 - y_2\|^{1/k}.$$

Observe that if Y is finite dimensional then (30) is always verified with $M=0$, $Q=B$. From Theorem 5.2 (vi) follows

COROLLARY 7.3. – *Assume that* $Y=\mathbf{R}^n$ *and that for some* $k \geqq 1$ *the convex cone spanned by* $G^k(x_0, y_0)$ *is equal to* Y. *Then the conclusions of Theorem* 7.2 *are valid.*

Proof. – By Theorem 2.1 it is enough to show that for some $\rho>0$ and all $(x, y) \in \operatorname{Graph}(G)$ near (x_0, y_0) and all small $h>0$

$$y + \rho h^k \mathring{B} \subset G(B_h(x)).$$

If we assume the contrary, then, by the proof of Theorem 3.2, there exist $(t_i, z_i) \to_G (x_0, y_0)$, $(x_i, y_i) \to_G (x_0, y_0)$, $\bar{y}_i \to y_0$, $h_i \to 0+$, $\bar{h}_i \to 0+$, $p_i \in Y^*$ of $\|p_i\|_{Y^*}=1$ and a function $\bar{o}: \mathbf{R}_+ \to \mathbf{R}_+$ such that $B_{\bar{h}_i}(x_i) \subset B_{h_i}(t_i)$, $\bar{y}_i \neq y_i$, $\lim_{h \to 0+} \bar{o}(h)/h = 0$ and

$$\forall (x, y) \in \operatorname{Graph}(G) \cap B_{h_i}(t_i) \times Y, \quad 0 \leqq \frac{1}{k} \langle p_i, y - y_i \rangle$$
$$+ \|y - y_i\| \bar{o}\left(\frac{\|y - y_i\|}{\|y_i - \bar{y}_i\|} \right) + \frac{1}{i} \|y_i - \bar{y}_i\|^{(k-1)/k} (d(x, x_i) + \|y - y_i\|) \quad (37)$$

and

$$\forall w \in \frac{G(B_{\bar{h}_i}(x_i)) - y_i}{\bar{h}_i^k},$$
$$0 \leqq \frac{1}{k} \langle p_i, w \rangle + \sqrt{i}\, \bar{o}\left(\frac{\|w\|}{\sqrt{i}} \right) + i^{-(k+1)/2k} (1 + \bar{h}_i^{k-1} \|w\|).$$

Fix $v \in G^k(x_0, y_0)$ and let $v_i \to v$ be such that $y_i + \bar{h}_i^k v_i \in G(B_{\bar{h}_i}(x_i))$. Thus we deduce from the last inequality

$$0 \leqq \frac{1}{k} \langle p_i, v_i \rangle + \sqrt{i}\, \bar{o}\left(\frac{\|v_i\|}{\sqrt{i}} \right) + i^{-(k+1)/2k} (1 + \bar{h}_i^{k-1} \|v_i\|)$$

Let p be a weak cluster point of $\{p_i\}$ (it exists because Y is reflexive). Then taking the limit in the above inequality we obtain $\langle p, v \rangle \geqq 0$ and since v is arbitrary

$$p \in (G^k(x_0, y_0))^+ = (\overline{co}\, G^k(x_0, y_0))^+ = \{0\}.$$

We show next that p can not be equal to zero. Fix $v \in G^{(1)}(x_i, y_i)$ and let $h_j \to 0+$, $v_j \to v$ be such that $y_i + h_j v_j \in G(B_{h_j}(x_i))$. Setting $y = y_i + h_j v_j$

in (37), dividing by h_j and taking the limit yields

$$0 \leqq \frac{1}{k} \langle p_i, v \rangle + \frac{1}{i} \| y_i - \bar{y}_i \|^{(k-1)/k} (1 + \| v \|).$$

Hence for a sequence $\varepsilon_i \to 0+$ we have

$$\forall v \in \overline{co} (\mathrm{G}^{(1)}(x_i, y_i) \cap \mathrm{MB}), \qquad \langle p_i, v \rangle \geqq -\varepsilon_i.$$

The end of the proof is similiar to that of Theorem 6.1. Let z, ρ, w_i, a_i, q_i be as in the proof of Theorem 6.1. Then

$$\langle p_i, z - q_i \rangle \geqq \langle p_i, \rho w_i \rangle + \langle p_i, a_i \rangle \geqq \rho \left(1 - \frac{1}{i}\right) - \varepsilon_i.$$

Consider subsequences $\{p_{i_j}\}$, $\{q_{i_j}\}$ such that $\{p_{i_j}\}$ converges weakly to p and $q_{i_j} \to q \in \mathrm{Q}$. Then the last inequality implies that $\langle p, z - q \rangle \geqq \rho$ which yelds that p can not be equal to zero and completes the proof. □

COROLLARY 7.4. – *Let* $y_0 \in \mathrm{G}(x_0)$. *Assume that* Y *is a Hilbert space and that there exists a (closed) subspace* $\mathrm{H} \subset \mathrm{Y}$ *of finite co-dimension*, $z \in \mathrm{H}$, $\rho > 0$, $\mathrm{M} > 0$ *such that for all* $(x, y) \in \mathrm{Graph}(\mathrm{G})$ *near* (x_0, y_0) *and all small* $h > 0$

$$z + \rho \mathrm{B_H} \subset \overline{co} (\mathrm{G}^{(1)}(x, y) \cap \mathrm{MB}).$$

If for some $k \geqq 1$, $0 \in \mathrm{Int}\, \overline{co}\, \mathrm{G}^k (x_0, y_0)$ *then we have the same conclusions as in Theorem* 7.2.

The proof is similar to the proof of Corollary 6.8 and is thus omitted.

8. TAYLOR EXPANSION AND THE INVERSE FUNCTION THEOREM

Consider a function f from a Banach space X to a uniformly smooth space Y. Let $x_0 \in \mathrm{X}$ and assume that for some integer $k \geqq 1$, $f \in \mathrm{C}^k$ at x_0. Then f can be approximated on a neighborhood of x_0 by its Taylor expansion:

Let $\dfrac{\partial^i f}{\partial x^i}(x)$ denote the $i-th$ derivative of f at x. Then there exists a function $o : \mathrm{R}_+ \to \mathrm{R}_+$ such that $\lim_{h \to 0+} \dfrac{o(h^k)}{h^k} = 0$ and for all x near x_0 and all $v \in \mathrm{B}$

$$\left\| f(x + hv) - f(x) - \sum_{i=1}^{k} \frac{h^i}{i!} \frac{\partial^i f}{\partial x^i}(x) v \ldots v \right\| \leqq o(h^k).$$

THEOREM 8.1. — *Assume that* $Y=\mathbf{R}^n$ *and*

$$0\in \operatorname{Int}\overline{co} \liminf_{x\to x_0,\ h\to 0+} \left\{ \sum_{i=1}^{k} \frac{h^{i-k}}{i!} \frac{\partial^i f}{\partial x^i}(x) v \ldots v \mid \|v\|\leq 1 \right\}.$$

Then there exists $L>0$ *such that for all* x *near* x_0 *and* y *near* y_0

$$\operatorname{dist}(x, f^{-1}(y)) \leq L \| f(x)-y \|^{1/k}.$$

The above is a consequence of Corollary 7.3 and the following.

LEMMA 8.2. — *The* $k-th$ *variation* $f^k(x_0, f(x_0))$ *is equal to*

$$\liminf_{x\to x_0,\ h\to 0+} \left\{ \sum_{i=1}^{k} \frac{h^{i-k}}{i!} \frac{\partial^i f}{\partial x^i}(x) v \ldots v \mid \|v\|\leq 1 \right\}.$$

Proof. — For all x near x_0, and all $v\in B$ and $h>0$ we have

$$f(x+hv)-f(x) = \sum_{i=1}^{k} \frac{h^i}{i!} \frac{\partial^i f}{\partial x^i}(x) v \ldots v + o(h^k, x)$$

where $\lim_{x\to x_0,\ h\to 0+} \frac{\|o(h^k, x)\|}{h^k}=0$. Hence the Hausdorff distance of $\left\{ \sum_{i=1}^{k} \frac{h^{i-k}}{i!} \frac{\partial^i f}{\partial x^i}(x) v \ldots v \mid \|v\|\leq 1 \right\}$ and $\frac{f(B_h(x))-f(x)}{h^k}$ is smaller that $\varepsilon(h)$, where $\lim_{h\to 0+} \varepsilon(h)=0$. The result follows from the definition of the $k-$th variation. □

Observe that Theorem 7.1 yields.

THEOREM 8.3. — *Assume that for some* $\rho>0$, $M>0$ *and for all* x *near* x_0 *and all small* $h>0$

$$\inf_{\|p\|_{Y^*}=1} \left\{ \langle p, w\rangle \mid w\in \left\{ \sum_{i=1}^{k} \frac{h^{i-k}}{i!} \frac{\partial^i f}{\partial x^i}(x) v \ldots v \mid \|v\|\leq 1 \right\}, \|w\|\leq M \right\} \geq \rho.$$

Then there exists $L>0$ *such that for all* x *near* x_0 *and* y *near* y_0

$$\operatorname{dist}(x, f^{-1}(y)) \leq L \| f(x)-y \|^{1/k}.$$

Remark. — In [17] we derived from the above theorem a second order sufficient condition for the existence of local Hölder inverse of f with the Hölder exponent equal to $\frac{1}{2}$, $(k=2)$.

9. AN IMPLICIT FUNCTION THEOREM

Consider Banach spaces X, P, Y and a continuous function $g : X \times P \to Y$. Assume that for some $(\bar{x}, \bar{p}) \in X \times P$

$$g(\bar{x}, \bar{p}) = 0$$

and $\frac{\partial g}{\partial x}(.,.)$ is continuous on a neighborhood of $(\bar{x}, \bar{p})$. We investigate here the sets

$$Z(p) = \{ x \in X : g(x, p) = 0 \}.$$

Since g is continuous, the (set-valued) map Z has a closed graph. Moreover

$$\bar{x} \in Z(\bar{p}).$$

The results of Section 6 yield the following implicit function theorem.

THEOREM 9.1. – *Assume that for some* $\rho > 0$

$$\rho B \subset \overline{\frac{\partial g}{\partial x}(\bar{x}, \bar{p}) B} \tag{38}$$

and let $0 < \varepsilon < 1$ *be given. Then for some* $\delta > 0$ *and all* $(x, p) \in B_\delta(\bar{x}) \times B_\delta(\bar{p})$ *there exists* $z(p) \in Z(p)$ (i. e, $g(z(p), p) = 0$) *such that*

$$\| x - z(p) \| \leqq \frac{1}{1-\varepsilon} \frac{1}{\rho} \| g(x, p) \|.$$

The above result extends [37], Theorem 2.2. The Banach open mapping theorem imply that (38) is verified (for some $\rho > 0$) when the derivative $\frac{\partial g}{\partial x}(\bar{x}, \bar{p})$ is surjective.

Proof. – Fix $\varepsilon' < \varepsilon < 1$ and let $\delta' > 0$ be such that for all $(x, p) \in B_{\delta'}(\bar{x}) \times B_{\delta'}(\bar{p})$

$$(1 - \varepsilon') \rho B \subset \overline{\frac{\partial g}{\partial x}(x, p) B}.$$

Let $0 < \delta < \frac{\delta'}{4}$ be so small that for all

$$(x, p) \in B_\delta(\bar{x}) \times B_\delta(\bar{p}), \qquad \| g(x, p) \| < \min \left\{ \frac{\delta'}{8}, \frac{\rho \delta' (1 - \varepsilon')}{4} \right\}.$$

For every $p \in B_\delta(\bar{p})$ define the (single-valued map) $G_p : X \to Y$ by

$$\forall x \in X, \quad G_p(x) = g(x, p).$$

Then G_p satisfies assumptions of Theorem 6.1 with ρ replaced by $(1-\varepsilon')\rho$ and $\varepsilon=\delta'$. Therefore for every $(x, y)\in B_{\delta'/4}(\bar{x})\times Y$ satisfying $\|G_p(x)\|\leqq \frac{\delta'}{4}$, $\|y-G_p(x)\|<\min\left\{\frac{\delta'}{8}, \frac{\delta'\rho(1-\varepsilon')}{4}\right\}$

$$\operatorname{dist}(x, G_p^{-1}(y))\leqq \frac{1}{\rho(1-\varepsilon')}\|G_p(x)-y\|<\frac{1}{\rho(1-\varepsilon)}\|G_p(x)-y\|$$

Applying the above with $y=0$ we obtain

$$\forall x\in B_\delta(\bar{x}),\quad \operatorname{dist}(x, Z(p))<\frac{1}{1-\varepsilon}\,\frac{1}{\rho}\|g(x, p)\|.$$

Since $p\in B_\delta(\bar{p})$ is arbitrary the proof is complete. □

Remark. – Theorem 9.1 can be proved for more general spaces P, Y and for a less regular function g. We do not do it here in order to simplify the presentation of the result. □

As an example of application consider a continuous function $f:\mathbf{R}^n\times\mathbf{R}^n\to\mathbf{R}^m$ and an implicit dynamical system.

$$f(x, x')=0 \tag{39}$$

An absolutely continuous function $x\in W^{1,1}(0,T)$, $T>0$ is called a trajectory of (39) if for almost all $t\in[0, T]$, $f(x(t), x'(t))=0$. A direct way to make the above system implicit is to replace (39) by the differential inclusion

$$x'\in F(x) \tag{40}$$

where the set-valued map $F:\mathbf{R}^n\to\mathbf{R}^n$ is given by

$$F(x)=\{\, v\in\mathbf{R}^n \mid f(x, v)=0\,\}.$$

An absolutely continuous function $x\in W^{1,1}(0, T)$, $T>0$ is called a trajectory of (40) if for almost all $t\in[0, T]$, $x'(t)\in F(x(t))$. It is clear that solutions of the implicit system and of the corresponding differential inclusion do coincide. Since f is continuous, F has a closed graph. If moreover for all $\bar{x}\in\mathbf{R}^n$ there exists $\varepsilon>0$ such that

$$\liminf_{\|v\|\to\infty}\ \inf_{x\in B_\varepsilon(\bar{x})}\|f(x, v)\|>0 \tag{41}$$

then F has compact images and is upper semicontinuous. When F is locally Lipschitz in the Hausdorff metric, then it inherits many properties of ODE. For instance, solution sets of the differential inclusion depend in a Lipschitz way on initial conditions [2] and the variational equation of ODE may be extended to differential inclusion ([20], [21]). The result below provides a sufficient condition for the local Lipschitz continuity of F (in the Hausdorff metric).

THEOREM 9.2. — *Assume that* $f \in C^1$ *and that* (41) *holds true for all* $x \in \mathbf{R}^n$. *If* $\bar{x} \in \mathbf{R}^n$ *is such that* $F(\bar{x}) \neq \emptyset$ *and for all* $v \in F(\bar{x})$ *the derivative* $\frac{\partial f}{\partial v}(\bar{x}, v)$ *is surjective, then* F *is locally Lipschitz at* $\bar{x}$.

Proof. — Fix $\bar{x}$ as above. By (41), $F(\bar{x})$ is a compact set. Hence the surjectivity of derivative yields that for some $\rho > 0$ and all $v \in F(\bar{x})$

$$\rho B \subset \frac{\partial f}{\partial v}(\bar{x}, v) B.$$

Since $f \in C^1$, by (41), there exist $L > 0$, $\bar{\delta} > 0$ such that for all $v \in F(B_{\bar{\delta}}(\bar{x}))$, $f(., v)$ is L-Lipschitz on $B_{\bar{\delta}}(\bar{x})$. Applying Theorem 9.1 to the function $g(x, p) := f(p, x)$ and using compactness of $F(\bar{x})$ we prove that for some $K > 0$, $0 < \delta < \bar{\delta}$ and all x, $y \in B_\delta(\bar{x})$, $v_x \in B_\delta(F(\bar{x})) \cap F(x)$ there exists $v_y \in F(y)$ satisfying

$$\| v_y - v_x \| \leq K \| f(y, v_x) \| \leq KL \| y - x \|. \tag{42}$$

On the other hand the upper semicontinuity of F yields that for some $0 < \delta' < \delta$ and all $x \in B_{\delta'}(\bar{x})$ we have $F(x) \subset B_\delta(F(\bar{x}))$. Therefore for all x, $y \in B_{\delta'}(\bar{x})$ and all $v_x \in F(x)$ there exists $v_y \in F(y)$ satisfying (42). This is equivalent to the local Lipschitz continuity of F at $\bar{x}$.

10. LIPSCHITZ BEHAVIOR OF CONTROLS

Let U be a separable metric space, E be a Banach space and $f : E \times U \to E$ be a continuous, differentiable in the first variable function. We assume that

(*a*) f is locally Lipschitz in the first variable uniformly on U, *i. e.* for every $x \in E$ there exist $L > 0$ and $\varepsilon > 0$ such that for all $u \in U$, $f(., u)$ is L-Lipschitz on $B_\varepsilon(x)$: *i. e.*

$$\forall x', \quad x'' \in B_\varepsilon(x), \qquad \| f(x', u) - f(x'', u) \| \leq L \| x' - x'' \|.$$

(*b*) For every $u \in U$ the derivative $\frac{\partial f}{\partial x}(., u)$ is continuous.

(*c*) For every $x \in E$ the set $f(x, U)$ is bounded.

For all $T > 0$ a (Lebesgue) measurable function $u : [0, T] \to U$ is called an admissible control. Let $\mathcal{U}_T$ denote the set of all admissible controls defined on the time interval [0, T]. Define a metric on $\mathcal{U}_T$ by setting

$$d_T(u, v) = \mu(\{ t \in [0, T] \mid u(t) \neq v(t) \})$$

where μ states for the Lebesgue measure. The space $(\mathcal{U}_T, d_T)$ is complete [13].

10.1. Finite dimensional control system

Let $E=\mathbf{R}^n$, $x_0 \in E$ and f, *U be as above. Consider the control system*

$$\left.\begin{aligned} x'=f(x, u(t)), \qquad u\in\mathcal{U}_T, T>0 \\ x(0)=x_0. \end{aligned}\right\} \tag{43}$$

An absolutely continuous function $x\in W^{1,1}(0, T)$ (the Sobolev space) is called trajectory of the control system (43) if $x(0)=x_0$ and there exists $u\in\mathcal{U}_T$ such that $x'(t)=f(x(t), u(t))$ *a. e.* in $[0, T]$.

For all $T>0$ the reachable set of the system (43) at time T is given by

$$R(T)=\{x(T)\mid x\in W^{1,1}(0,T) \text{ is a trajectory of } (43)\}.$$

Let $z\in W^{1,1}(0, T)$ be a given trajectory and $\bar{u}\in\mathcal{U}_T$ be a corresponding control. We provide here a sufficient condition for

$$z(T)\in \operatorname{Int} R(T)$$

and for regularity of the "inverse".

Consider the linear control system

$$\begin{aligned} w'(t)=\frac{\partial f}{\partial x}(z(t),\bar{u}(t))\,w(t)+y(t), \quad y(t)\in\overline{co}\, f(z(t),U)-f(z(t),\bar{u}(t)) \\ w(0)=0 \end{aligned} \tag{44}$$

and define the corresponding reachable set

$$R^L(T)=\{w(T)\mid w\in W^{1,1}(0,T) \text{ is a trajectory of } (44)\}.$$

For all $s\in[0, T]$, let $S_{\bar{u}}(\,.\,; s)$ denote the solution matrix of the system

$$\begin{aligned} Z'(t)=\frac{\partial f}{\partial x}(z(t),\bar{u}(t))\,Z(t), \qquad t\in[s,T] \\ Z(s)=\mathbf{1} \end{aligned}$$

where $\mathbf{1}$ states for the identity. Then

$$R^L(T)=\left\{\int_0^T S_{\bar{u}}(T; s)\,y(s)\,ds \,\middle|\, y(s)\in\overline{co}\, f(z(s),U)-f(z(s),\bar{u}(s))\right\}.$$

For every $u\in\mathcal{U}_T$ we denote by x_u the solution of (43) (when it is defined on the whole time interval $[0, T]$) corresponding to the control u.

THEOREM 10.1. – *Assume that*

$$0\in \operatorname{Int} R^L(T) \tag{45}$$

Then $z(T)\in \operatorname{Int} R(T)$ *and there exist* $\varepsilon>0$, $L>0$ *such that for every control* $u\in\mathcal{U}_T$ *satisfying* $d_T(u, \bar{u})\leq\varepsilon$ *and every* $b\in B_\varepsilon(z(T))$ *we can find a control* $\hat{u}\in\mathcal{U}_T$ *with*

$$x_{\hat{u}}(T)=b; \qquad \mu(\{t\in[0,T]\mid \hat{u}(t)\neq u(t)\})\leq L\,\|b-x_u(T)\|$$

In particular for all $b \in B_\varepsilon(z(T))$ *there exists a control* $u \in \mathcal{U}_T$ *such that*

$$x_u(t)=b, \qquad \mu(\{t\in[0,T]\,|\,u(t)\neq\bar{u}(t)\})\leqq L\|b-z(T)\|.$$

Remark. – The fact that from (45) follows $z(T)\in \operatorname{Int} R(T)$ is well known. The second part of Theorem 10.1 providing an estimation of controls is a new results. We also observe that ε and L in the above claim can be estimated from the data of the problem. □

Proof. – Replacing t by $\frac{t}{T}$ we may assume that $T=1$. Set $\mathcal{U}=\mathcal{U}_1$. From the Gronwall inequality follows that for some $\delta>0$, the map $\varphi(u)=x_u$ from $B_\delta(\bar{u})$ to $C(0,1;E)$ is single-valued and Lipschitz continuous. For all $u\in B_\delta(\bar{u})$ and $s\in[0,1]$ let $S_u(.\,;s)$ denote the solution matrix of the linear system

$$Z'(t)=\frac{\partial f}{\partial x}(x_u(t),u(t))Z(t), \qquad t\in[s,1]$$
$$Z(s)=\mathbf{1}$$

Fix $u\in B_{\delta/2}(\bar{u})$, $v\in U$ and let $0<t_0<1$ be such that $x_u'(t_0)=f(x_u(t_0), u(t_0))$. (The set of such points t_0 is of full measure in [0, 1].) For all small $h>0$ consider controls

$$u_h(t)=\begin{cases} v, & t_0-h<t\leqq t_0 \\ u(t), & \text{otherwise}\end{cases} \tag{46}$$

and let x_h denote the solution of (43) corresponding to u_h. Controls (46) are needle perturbations of u and it is well known that

$$\lim_{h\to 0+}\frac{x_h(1)-x_u(1)}{h}=S_u(1\,;t_0)(f(x_u(t_0),v)-f(x_u(t_0),u(t_0))). \tag{47}$$

Set $V_u(t)=f(x_u(t),U)-f(x_u(t),u(t))$ and define the Lipschitz continuous map $G:B_\delta(\bar{u})\to E$

$$G(u)=\varphi(u)(1)=x_u(1).$$

Then, by (47), for every arbitrary but fixed $u\in B_{\delta/2}(\bar{u})$, for almost all $t_0\in[0,1]$ and all $v\in V_u(t_0)$, $S_u(1\,;t_0)v\in G^{(1)}(u, x_u(1))$. Let M be the Lipschitz constant of G. Hence $G^{(1)}(u,x_u(1))\subset MB$ and we proved that for almost every $t_0\in[0,1]$, and every

$$y\in\overline{co}\,V_u(t_0), \qquad S_u(1\,;t_0)y\in\overline{co}(G^{(1)}(u,x_u(1))\cap MB).$$

From the mean value theorem follows that for all measurable selection $y(t)\in\overline{co}\,V_u(t)$

$$\int_0^1 S_u(1\,;t)y(t)\,dt\in\overline{co}(G^{(1)}(u,x_u(1))\cap MB). \tag{48}$$

By (45) there exists $\rho > 0$ such that

$$\rho \mathrm{B} \subset \left\{ \int_0^1 \mathrm{S}_{\bar{u}}(1\,;\,t)\, y(t)\, dt \,|\, y(t) \in \overline{co}\, \mathrm{V}_{\bar{u}}(t) \right\}. \tag{49}$$

On the other hand Gronwall's inequality and the assumptions (*a*), (*b*) imply that $\{ \mathrm{S}_u(1\,;\,.) \}$ converges uniformly to $\mathrm{S}_{\bar{u}}(1\,;\,.)$ when $u \to \bar{u}$ and from (*a*), (*c*) and continuity of f we deduce that

$$\lim_{u \to \bar{u}} \int_0^1 \mathscr{H}(\overline{co}\, \mathrm{V}_{\bar{u}}(t),\, \overline{co}\, \mathrm{V}_u(t))\, dt = 0$$

where $\mathscr{H}$ states for the Hausdorff distance. Since the right-hand side of (48) is convex, this yields that for some $0 < \bar{\delta} < \frac{1}{2}\delta$ and all $u \in \mathrm{B}_{\bar{\delta}}(\bar{u})$

$$\frac{\rho}{2}\mathrm{B} \subset \left\{ \int_0^1 \mathrm{S}_u(1\,;\,t)\, y(t)\, dt \,|\, y(t) \in \overline{co}\, \mathrm{V}_u(t) \right\} \subset \overline{co}\,(\mathrm{G}^{(1)}(u, x_u(1)) \cap \mathrm{MB}) \tag{50}$$

Theorem 6.5 then and (50), (48) end the proof.

10.2. Infinite dimensional control system

We assume here that E is a Banach space the norm of which is Gâteaux differentiable away from zero. Let $\{ \mathrm{S}(t) \}_{t \geqq 0}$ be a strongly continuous semigroup of linear operators from E to E and A be its infinitesimal generator, $x_0 \in \mathrm{E}$. Consider the control system

$$\left. \begin{aligned} x'(t) = \mathrm{A}\, x(t) + f(x(t), u(t)), \qquad u \in \mathscr{U}_\mathrm{T}, \quad \mathrm{T} > 0 \\ x(0) = x_0 \end{aligned} \right\} \tag{51}$$

Recall that a continuous function $x : [0, \mathrm{T}] \to \mathrm{E}$ is called a mild trajectory of (51) corresponding to the control $u \in \mathscr{U}_\mathrm{T}$ if for all $0 \leqq t \leqq \mathrm{T}$

$$x(t) = \mathrm{S}(t)\, x_0 + \int_0^t \mathrm{S}(t-s)\, f(x(s), u(s))\, ds.$$

As before we define the reachable set of (51) at time $\mathrm{T} > 0$ by

$$\mathrm{R}(\mathrm{T}) = \{ x(\mathrm{T}) \,|\, x \in \mathrm{C}(0, \mathrm{T}\,;\, \mathrm{E}) \text{ is a mild trajectory of (51)} \}.$$

Let $(z, \bar{u})$ be a trajectory-control pair of (51) on [0, T]. We study the same question as in the previous section, *i. e.* sufficient condition for

$$z(\mathrm{T}) \in \mathrm{Int}\, \mathrm{R}(\mathrm{T})$$

and for regularity of the "inverse". Consider the linear control system

$$\begin{gathered} w'(t)=\mathrm{A}\,w(t)+\frac{\partial f}{\partial x}(z(t),\bar{u}(t))\,w(t),\\ y(t)\in\overline{co}\,f(z(t),\mathrm{U})-f(z(t),\bar{u}(t))\\ w(0)=0 \end{gathered} \tag{52}$$

Let $\mathrm{S}_{\bar{u}}(t;s)$ denote the solution operator of the equation

$$\mathrm{Z}'(t)=\mathrm{AZ}(t)+\frac{\partial f}{\partial x}(z(t),\bar{u}(t))\,\mathrm{Z}(t) \tag{53}$$

That is the only strongly continuous solution of the operator equation

$$\forall p\in\mathrm{E},$$

$$\mathrm{S}_{\bar{u}}(t;s)p=\mathrm{S}(t-s)p+\int_s^t \mathrm{S}(t-\sigma)\frac{\partial f}{\partial x}f(z(\sigma),\bar{u}(\sigma))\,\mathrm{S}_{\bar{u}}(\sigma;s)p\,d\sigma$$

where $0\leqq s\leqq t\leqq 1$. Then the reachable set $\mathrm{R}^{\mathrm{L}}(\mathrm{T})$ of (52) by the mild trajectories at time T is given by

$$\mathrm{R}^{\mathrm{L}}(\mathrm{T})=\left\{\int_0^{\mathrm{T}}\mathrm{S}_{\bar{u}}(\mathrm{T};s)\,y(s)\,ds\,\middle|\,y(s)\in\overline{co}f(z(s),\mathrm{U})-f(z(s),\bar{u}(s))\right\}.$$

For every $u\in\mathscr{U}_{\mathrm{T}}$ denote by x_u the mild solution of (51) (when it is defined on the whole time interval [0, T]) corresponding to the control u.

THEOREM 10.2. – *Assume that* $0\in\operatorname{Int}\mathrm{R}^{\mathrm{L}}(\mathrm{T})$. *Then* $z(\mathrm{T})\in\operatorname{Int}\mathrm{R}(\mathrm{T})$ *and there exist* $\varepsilon>0$, $\mathrm{L}>0$ *such that for every control* $u\in\mathscr{U}_{\mathrm{T}}$ *satisfying* $d_{\mathrm{T}}(u,\bar{u})\leqq\varepsilon$ *and all* $b\in\mathrm{B}_{\varepsilon}(z(\mathrm{T}))$ *there exists* $\hat{u}\in\mathscr{U}_{\mathrm{T}}$ *with*

$$x_{\hat{u}}(\mathrm{T})=b;\qquad \mu(\{t\in[0,\mathrm{T}]\,|\,u(t)\neq\hat{u}(t)\})\leqq\mathrm{L}\,\|b-x_u(\mathrm{T})\|.$$

Proof. – The proof is analogous to the proof of Theorem 10.1 so we only sketch it. We may assume again that $\mathrm{T}=1$ and applying the Gronwall inequality, we can find $\delta>0$ such that the map $\varphi:\mathrm{B}_{\delta}(\bar{u})\to\mathrm{C}(0,1;\mathrm{E})$ associating with every $u\in\mathrm{B}_{\delta}(\bar{u})$ the mild solution x_u of (51) is Lipschitz continuous. For all $u\in\mathrm{B}_{\delta}(\bar{u})$, let $\mathrm{S}_u(t;s)$ denote the solution operator of the equation

$$\mathrm{Z}'(t)=\mathrm{AZ}(t)+\frac{\partial f}{\partial x}(x_u(t),u(t))\,\mathrm{Z}(t). \tag{54}$$

Define the continuous map $\mathrm{G}:\mathrm{B}_{\delta}(\bar{u})\to\mathrm{E}$ by $\mathrm{G}(u)=x_u(1)$ and consider again the needle perturbations of controls (46) and the corresponding trajectories x_h of (51). Then we obtain (47) for all fixed u near $\bar{u}$ and all t_0 from the set of Lebesgue points of $f(x_u(.),u(.))$ of full measure (*see* for example [14]). Let $\mathrm{V}_u(t)$ be defined as in the proof of Theorem 10.1. By the same arguments for all u near $\bar{u}$ and every measurable selection

$y(t) \in \overline{co}\, V_u(t)$ the inclusion (48) holds true. The assumptions of theorem yield (49) for some $\rho > 0$. The Gronwall inequality imply that (50) holds true for all u near $\bar{u}$ by the same reasons as in the proof of theorem 10.1. Theorem 6.5 ends the proof.

11. A MULTIPLIER RULE FOR INFINITE DIMENSIONAL PROBLEMS

We study here necessary conditions satisfied by optimal solutions to the abstract optimization problem

$$\text{minimize}\{ J(u) \mid u \in \mathcal{U}, G(u) \in K \} \tag{55}$$

where

$\mathcal{U}$, is a complete metric space
J, is a locally Lipschitz function from $\mathcal{U}$ to **R**
G, is a locally Lipschitz function from $\mathcal{U}$ to a Banach space X
K, is a closed subset of X.

We denote by $\| . \|$ the norm of X.

Recall that the contingent cone to K at $x \in K$ is defined by

$$T_K(x) = \limsup_{h \to 0+} \frac{K - x}{h}$$

and Clarke's tangent cone to K at $x \in K$ by

$$C_K(x) = \liminf_{x' \to_K x,\, h \to 0+} \frac{K - x'}{h}$$

where $\to_K$ denotes the convergence in K.

The normal cone to K at x is the negative polar of $C_K(x)$:

$$N_K(x) = C_K(x)^- = \{ p \in Y^* \mid \forall\, w \in C_K(x), \langle p, w \rangle \leqq 0 \}.$$

The cone $C_K(x)$ is convex and is contained in $T_K(x)$ ([9], [3]).

DEFINITION 11.1. – *We say that the set* K *is sleek near* $x_0 \in K$ *if there exists a neighborhood* $\mathcal{N}$ *of* x_0 *in* K *such that for every* $x \in \mathcal{N}$

$$T_K(x) = \liminf_{x' \to_K x} T_K(x')$$

In this case for every $x \in \mathcal{N}$, $T_K(x) = C_K(x)$ [3].

In particular K is sleek when it is convex or when it is a C^1-manifold.

THEOREM 11.2. – *Assume that* u_0 *solves the problem* (55) *and that* K *is sleek near* $G(u_0)$. *Further assume that for some compact* $Q \subset X$, $\rho > 0$, $\varepsilon > 0$

and closed convex sets $C(u) \subset (J, G)^{(1)}(u, J(u), G(u))$ *containing zero we have*

$$\forall (u,k) \in B_\varepsilon(u_0) \times (K \cap B_\varepsilon(G(u_0))), \quad \rho B \subset \pi(C(u)) - T_K(x) + Q \quad (56)$$

where π *denotes the canonical projection of* $\mathbf{R} \times X$ *onto* X. *Then there exist*

$$\lambda \geqq 0, \quad p \in N_K(G(u_0)), \quad (\lambda, p) \neq 0 \quad (57)$$

such that

$$\forall (j,g) \in \liminf_{u \to u_0} C(u), \quad \lambda j + \langle p, g \rangle \geqq 0. \quad (58)$$

Remark.

(*a*) Observe that when X is a finite dimensional space, then the condition (56) is always satisfied with Q equal to the unit ball and $\rho = 1$.

(*b*) It is possible to prove a similar theorem without assuming that K is sleek, by using closed convex subcones of $T_K(x)$ and their lower limits. This will lead however to somewhat "heavy" formulas. The sleekness hypothezis allows to avoid such misbehavior of nonsmooth sets of constraints. □

Proof. – Consider the set-valued map $P: \mathcal{U} \to \mathbf{R} \times X$ defined by

$$P(u) = (J(u) + \mathbf{R}_+, G(u) - K).$$

We first verify that *Graph* (P) is closed. Indeed let $(u_n, q_n) \in \text{Graph}(P)$ be a sequence converging to some (u, q). Then for some $r_n \geqq 0$, $k_n \in K$, $q_n = (J(u_n) + r_n, G(u_n) - k_n)$. Since J, G are continuous $(r_n, -k_n) \to q - (J(u), G(u))$ and thus $\{ r_n \}$ and $\{ k_n \}$ are converging to some $(r, k) \in \mathbf{R}_+ \times K$. From now on we write $(J, G)^{(1)}(u)$ for $(J, G)^{(1)}(u, J(u), G(u))$. It is not difficult to show that for all $u \in \mathcal{U}$, $r \geqq 0$, $k \in K$

$$(J, G)^{(1)}(u) - \mathbf{R}_- \times T_K(k) \subset P^{(1)}(u, J(u) + r, G(u) - k) \quad (59)$$

Hence for all $u \in \mathcal{U}$, $r \geqq 0$, $k \in K$

$$A(u,k) := C(u) - \mathbf{R}_- \times T_K(k) \subset P^{(1)}(u, J(u) + r, G(u) - k).$$

Since J is locally Lipschitz there exists $M > 0$ such that for all $u \in \mathcal{U}$ near u_0 and for every $(v, w) \in C(u) \subset \mathbf{R} \times X$ we have $\|v\| \leqq M$. This and the assumption (56) yields that for all $u \in \mathcal{U}$ near u_0 and all $k \in K$ near $G(u_0)$

$$[-1, 1] \times \rho B \subset A(u,k) + [-M-1, M+1] \times Q$$

The sets $\overline{A(u,k)}$ being convex closed, we may use Theorem 6.4 with G equal to P. Since u_0 is an optimal solution $(J(u_0), 0)$ is a boundary point of Im P (image of P). Therefore the second statement of Theorem 6.4 holds true. Thus for some $\lambda \in \mathbf{R}$, $p \in X^*$ not both equal to zero

$$\forall (j,g) \in \liminf_{u \to u_0} C(u), \quad r \geqq 0, \quad v \in \liminf_{k \to_K G(u_0)} T_K(k),$$

$$\lambda(j+r)+\langle p, g-v \rangle \geqq 0.$$

Using that K is sleek we deduce that

$$\liminf_{k \to_K G(u_0)} T_K(k) = T_K(G(u_0)) = C_K(G(u_0)).$$

Setting $(j, g)=0$ we obtain (57) and setting $r=0$, $v=0$ we get (58). □

When the norm of X is differentiable, then the sleekness hypothezis on K may be omitted and a stronger result can be proved:

THEOREM 11.3. – *Assume that the norm of* X *is Gâteaux differentiable away from zero and let* $u_0 \in \mathcal{U}$ *be an optimal solution of* (55). *Further assume that for some compact set* $Q \subset X$, $\rho>0$, $\gamma>0$, $\varepsilon>0$

$$\forall (u, x) \in B_\varepsilon(u_0) \times (K \cap B_\varepsilon(G(u_0))),$$
$$\rho B \subset \overline{co}(G^{(1)}(u, G(u)) - T_K(x) \cap \gamma B) + Q.$$

Then there exist (λ, p) *as in* (57) *such that*

$$\forall (j, g) \in \liminf_{u \to u_0} \overline{co}(J, G)^{(1)}(u, J(u), G(u)), \quad \lambda j + \langle p, g \rangle \geqq 0. \tag{60}$$

Observe that if $Q \subset X$ is compact, so is $\overline{co}\, Q$. Therefore we may always assume in Theorem 11.3 that Q is convex. This and the separation theorem immediatly yield

THEOREM 11.4. – *Let* X, u_0 *be as in Theorem* 11.3. *Further assume that there exist subsets* $A(u) \subset \overline{co}\, G^{(1)}(u, G(u))$, *such that the map* $u \to A(u)$ *is continuous at* u_0. *If for some compact set* $Q \subset X$, $\rho>0$, $\gamma>0$ *and all* $x \in K$ *near* $G(u_0)$

$$\rho B \subset \overline{co}(A(u_0) - T_K(x) \cap \gamma B) + Q$$

then the same assertions as in Theorem 11.3 *are valid.*

COROLLARY 11.5. – *In Theorem* 11.4 *assume that* $J = \varphi \circ G$, *where* $\varphi : X \to \mathbf{R}$ *is* C^1 *at* $G(u_0)$. *Then there exist* λ, p *as in* (57) *such that*

$$\forall w \in A(u_0), \quad \langle \lambda \varphi'(G(u_0)) + p, w \rangle \geqq 0.$$

Proof. – Let $(J, G)^{(1)}(u)$ and the set-valued map P be defined as in the proof of Theorem 11.2. We already know that Graph (P) is closed and that (59) holds true for all $u \in \mathcal{U}$, $r \geqq 0$, $k \in K$. It is also not difficult to check that

$$\mathbf{R}_+ \times (-C_K(G(u_0))) \in P^1(u_0, J(u_0), 0). \tag{61}$$

Since J and G are locally Lipschitz, there exists $M>0$ such that for all u near u_0, $(J, G)^{(1)}(u) \subset MB$. On the other hand for every $w \in G^{(1)}(u, G(u))$ there exists $v \in \mathbf{R}$ such that $(v, w) \in (J, G)^{(1)}(u)$. Indeed let $h_i \to 0+$, $w_i \to w$, $u_i \in B_{h_i}(u)$ be such that $G(u) + h_i w_i = G(u_i)$. Since J is Lipschitz at u_0 the sequence $\left\{\frac{J(u_i)-J(u)}{h_i}\right\}$ is bounded. Hence taking a subsequence and keeping the same notations we may assume that for some $v\in\mathbf{R}$

$$v_i := \frac{J(u_i)-J(u)}{h_i} \to v.$$

Thus $(J(u)+h_i v_i, G(u)+h_i w_i)\in(J, G)(B_{h_i}(u))$ and therefore $(v, w)\in(J, G)^{(1)}(u)$. This and the assumption of theorem yield that for all $u\in\mathscr{U}$ near u_0, for all $r\geqq 0$ and all $k\in K$ near $G(u_0)$

$$\begin{aligned}[-1,1]\times\rho B &\subset \overline{co}((J,G)^{(1)}(u)-[-1,0]T_K(x)\cap\gamma B)\\ &\qquad\qquad +[-M-2,M+2]\times Q\\ &\subset \overline{co}(P^{(1)}(u, J(u)+r, G(u)-k)\cap\sqrt{M^2+\gamma^2+1}\,B)+[-M-2,M+2]\times Q\end{aligned}$$

Since u_0 is an optimal solution $(J(u_0), 0)$ is a boundary point of Im P. Hence by Theorem 6.7 there exist $\lambda\in\mathbf{R}$, $p\in X^*$ not both equal to zero such that

$$\forall(v, w)\in \liminf_{(u,t,z)\to_P(u_0,J(u_0),0)} \overline{co}(P^{(1)}(u,t,z)\cap\sqrt{M^2+\gamma^2+1}\,B) + P^1(u_0, J(u_0), 0),\qquad \lambda v+\langle p, w\rangle\geqq 0$$

Then from (61) we deduce (57). On the other hand from (59)

$$\liminf_{u\to u_0}\overline{co}(J,G)^{(1)}(u)\subset \liminf_{(u,t,z)\to_P(u_0,J(u_0),0)} \overline{co}(P^{(1)}(u,t,z)\cap\sqrt{M^2+\gamma^2+1}\,B)$$

and we get (60). □

We apply the above results to derive necessary conditions for optimality for two infinite dimensional control problems:

11.1. A Semilinear control problem with end point constraints

Let E be a Banach space whose norm is Gâteaux differentiable away from zero. Consider a C^1-function $\varphi: E\times E\to\mathbf{R}$, $T>0$ and closed subsets $K_0, K_T\subset E$. We study the optimal control problem

$$\text{minimize}\,\varphi(x(0), x(T)) \tag{62}$$

over solutions of the system

$$\left.\begin{aligned} x'(t)=A\,x(t)+f(x(t), u(t)),\qquad u\in\mathscr{U}_T\\ x(0)\in K_0,\qquad x(T)\in K_T\end{aligned}\right\}\tag{63}$$

where A, f, $\mathscr{U}_T$ have the same meaning as in the Section 10 and satisfy the same assumptions. Let $S_{\bar u}(\,.\,;\,.\,)$ be the solution operator of (53).

THEOREM 11.6. – *Let* $(z, \bar{u})$ *be an optimal trajectory-control pair of the problem* (62)-(63) *and let* $R^L(T)$ *denotes the reachable set of the linearized system* (52) *at time* T. *Assume that for some* $\rho>0$, $\gamma>0$, $\varepsilon>0$ *and a compact set* $Q \subset X$

$$\forall x \in K_T \cap B_\varepsilon(z(T)), \quad \rho B \subset R^L(T) - \overline{co}(T_{K_T}(x) \cap \gamma B) + Q \tag{64}$$

Then there exist $\lambda \geqq 0$, $\xi_0 \in N_{K_0}(z(0))$, $\xi_T \in N_K(z(T))$ *not vanishing simultaneously such that the function*

$$p(t) = S_{\bar{u}}(T; t)^* \left(-\lambda \frac{\partial \varphi}{\partial x_2}(z(0), z(T)) - \xi_T \right) \tag{65}$$

satisfies the maximum principle

$$\langle p(t), f(z(t), \bar{u}(t)) \rangle = \max_{u \in U} \langle p(t), f(z(t), u) \rangle \quad \text{a.e. } in\ [0, T] \tag{66}$$

and the transversality condition

$$(p(0), -p(T)) = \lambda \varphi'(z(0), z(T)) + (\xi_0, \xi_T). \tag{67}$$

Proof. – It is not restrictive to assume that $T=1$. Set $\mathcal{U} = \mathcal{U}_1$. We apply results of Section 10. Define the locally Lipschitz map $G: E \times \mathcal{U} \to E \times E$ by

$$\forall x_0 \in E, \quad \forall u \in \mathcal{U}, \qquad G(x_0, u) = (x_0, x_u(1))$$

where x_u denotes the solution of (51) corresponding to the control u. By the Gronwall inequality G is locally Lipschitz. Set $V_{x_0, u}(t) = f(x_u(t), U) - f(x_u(t), u(t))$ and

$$R^L_{x_0, u}(1) = \left\{ \int_0^1 S_{x_0, u}(1; t) y(t)\, dt \,\middle|\, y(t) \in \overline{co}\, V_{x_0, u}(t) \right\}$$

where $S_{x_0; u}(\cdot\,;\cdot)$ denotes the solution operator of (54). From Section 10.2 we known that for all $u \in \mathcal{U}$ near u_0

$$\{0\} \times R^L_{x_0, u}(1) \subset \overline{co}(G^{(1)}(x_0, u, G(x_0, u))).$$

Differentiating with respect to the initial condition, we obtain easily that for all $x_0 \in E$, $w \in E$, $u \in \mathcal{U}$

$$\lim_{h \to 0+} \frac{G(x_0 + hw, u) - G(x_0, u)}{h} = S_{x_0, u}(1; 0) w.$$

Therefore for all $u \in \mathcal{U}$ near u_0

$$A(x_0, u) := \frac{1}{2} \{ (w, S_{x_0, u}(1; 0) w) \mid \|w\| \leqq 1 \}$$
$$+ \{0\} \times \frac{1}{2} R^L_{x_0, u}(1) \subset \overline{co}\, G^{(1)}(x_0, u, G(x_0, u)).$$

On the other hand, by Gronwall's lemma and assumptions (*a*)-(*c*) from Section 10, the map $(x_0, u) \to A(x_0, u)$ is continuous in the Hausdorff metric. Let $\delta > 0$ be so small that for all $\|w\| \leqq \delta$, $\|S_{\bar{u}}(1, 0) w\| \leqq \frac{\rho}{2}$. Thus, from the assumption (64), for all $(y_0, y_1) \in K_0 \times K_1$ near $(z(0), z(1))$

$$\frac{\delta}{2} \times \frac{\rho}{4} B \subset A(x_0, \bar{u}) - \overline{co}(T_{K_0}(y_0) \cap B) \times \overline{co}\left(T_{K_1}(y_1) \cap \frac{\gamma}{2} B\right)$$
$$+ \{0\} \times \frac{1}{2} Q \subset \overline{co}\,(G^{(1)}(z(0), \bar{u}, z(1))$$
$$- (T_{K_0}(y_0) \cap B)(T_{K_1}(y_1) \cap \gamma B)) + \{0\} \times \frac{1}{2} Q.$$

Hence, applying Corollary 11.5 we deduce that for some $\lambda \geqq 0$, $(\xi_0, \xi_1) \in N_{K_0 \times K_1}(z(0), z(1)) = N_{K_0}(z(0)) \times N_{K_1}(z(1))$ not all equal to zero

$$\left.\begin{array}{c} \forall w \in B, \quad r \in R^L(1), \\ \langle \lambda \varphi'(z(0), z(1)) + (\xi_0, \xi_1), (w, S_{\bar{u}}(1; 0) w + r) \rangle \geqq 0 \end{array}\right\} \qquad (68)$$

Define p by (65) with $T = 1$. Setting $r = 0$ in the above we get

$$\forall w \in B, \quad \left\langle \lambda \frac{\partial \varphi}{\partial x_1}(z(0), z(1)) + \xi_0 - p(0), w \right\rangle \geqq 0$$

and therefore $p(0) = \lambda \dfrac{\partial \varphi}{\partial x_1}(z(0), z(1)) + \xi_0$. Setting $w = 0$ in (68) yields that for every measurable selection $v(t) \in \overline{co} f(z(t), U) - f(z(t), \bar{u}(t))$,

$$\left\langle \lambda \frac{\partial \varphi}{\partial x_2}(z(0), z(1)) + \xi_1, \int_0^1 S_{\bar{u}}(1; t) v(t)\, dt \right\rangle = \int_0^1 \langle -p(t), v(t) \rangle\, dt \geqq 0$$

and the maximum principle (66) follows. From the definition of p we get the transversality condition (67). □

11.2 Optimal control of a problem with state constraints

Let Ω be an open bounded subset of $\mathbf{R}^n$, $(n \leqq 3)$ with C^2 boundary Γ, X be a Banach space and $L : C_0(\Omega) \to X$ be a C^1-mapping. Set $Y = H^2(\Omega) \cap H_0^1(\Omega)$ and consider sleek closed sets $K \subset L^2(\Omega)$, $D \subset X$ and a continuously differentiable function $J : C_0(\Omega) \times L^2(\Omega) \to \mathbf{R}$. We study the problem

$$\text{minimize } J(y, u)$$

over the pair $(y, u) \in Y \times K$ satisfying

$$A\,y + \varphi(y) = u \quad \text{in } \Omega$$
$$y = 0 \quad \text{on } \Gamma$$
$$L(y) \in D$$

where

$$A\,y = -\sum_{i,\,j=1}^{n} \partial_{x_j}(\alpha_{i,\,j}(x)\,\partial_{x_i} y) + a_0(x)\,y$$

and

$$a_0 \in L^{\infty}(\Omega), \quad a_0(x) \geqq 0 \quad \text{for } a.e.\ x \in \Omega, \quad a_{i,\,j} \text{ is Lipschitz on } \Omega$$
$$\forall \xi \in \mathbf{R}^n, \quad x \in \Omega, \quad \sum_{i,\,j=1}^{n} a_{i,\,j}(x)\,\xi_i \xi_j \geqq \alpha_0 \|\xi\|^2 \text{ for some } \alpha_0 > 0$$
$$\varphi \in C^1 \text{ is a nondecreasing function,}$$

A similar problem was considered in [7] with convex sets K, D and an arbitrary C^1-function φ.

THEOREM 11.7. – *Let $(\bar{y}, \bar{u})$ be an optimal solution of the above problem and B_X denote the closed unit ball in* X. *Further assume that for some* $\rho > 0$ *and a compact* $Q \subset X$, *we have*

$$\forall\, d \in D \textit{ near } L(\bar{y}), \quad \rho\, B_X \subset T_D(d) + Q. \tag{69}$$

Then there exist $\lambda \geqq 0$, $\mu \in N_D(\bar{y})$, $p \in W^{1,\,s}(\Omega)$, $s \in \left]0, \frac{n}{n-1}\right[$ *not all equal to zero, such that*

$$A^* p + \varphi'(\bar{y})^* p = \lambda \frac{\partial J}{\partial y}(\bar{y}, \bar{u}) + L'(\bar{y})^* \mu \tag{70}$$

$$-\lambda \frac{\partial J}{\partial u}(\bar{y}, \bar{u}) - p \in N_K(\bar{u}) \tag{71}$$

Proof. – Replacing Q by $\overline{co}\,Q$ we may assume that Q is convex. Define $A_1 : Y \to L^2(\Omega)$, $J_1 : Y \to \mathbf{R}$, $G : Y \to X \times L^2(\Omega)$ by

$$A_1(y) = A\,y + \varphi(y), \qquad J_1(y) = J(y, A_1(y)), \quad G(y) = (L(y), A_1(y))$$

and set $\mathscr{K} = D \times K$. Then, K and D being sleek,

$$\forall (d, k) \in D \times K, \quad T_{\mathscr{K}}(d, k) = T_D(d) \times T_K(k) \tag{72}$$

and our problem reduces to

$$\text{minimize } \{ J_1(y) \mid y \in Y,\ G(y) \in \mathscr{K} \}.$$

A direct computation yields that for all $y \in Y$ and all $w \in B_Y$

$$(J_1'(y)(w, A\,w + \varphi'(y)\,w),\ L'(y)\,w,\ A\,w + \varphi'(y)\,w) \subset (J, G)^{(1)}(y, J(y), G(y)).$$

Moreover, since φ is nondecreasing, $\forall x \in \Omega$, $\varphi'(\bar{y}(x)) \geqq 0$. Hence from [35] follows that for some $\varepsilon > 0$

$$\varepsilon B_{L^2(\Omega)} \subset (A + \varphi'(\bar{y})) B_Y.$$

This and the assumption (69) imply that for some $\bar{\rho} > 0$ and all $(d, k) \in \mathscr{K}$ near $(L(\bar{y}), A_1(\bar{y}))$

$$\bar{\rho} B \subset cl\{(L'(\bar{y}) w, A w + \varphi'(\bar{y}) w) \mid \|w\| \leqq 1\} - T_D(d) \times T_K(k) - Q \times \{0\}$$

where cl denotes the closure. Since the map

$$y \to Z(y) := cl\{(L'(y) w, A w + \varphi'(y) w) \mid \|w\| \leqq 1\} \subset G^{(1)}(y, G(y))$$

is continuous in the Hausdorff metric, using the separation theorem, (72) and convexity of tangent cones $T_D(d)$, $T_K(k)$, we prove that for all $k \in \mathscr{K}$ near $(L(\bar{y}), A_1(\bar{y}))$ and all $y \in Y$ near $\bar{y}$

$$\frac{\bar{\rho}}{2} B \subset Z(y) - T_{\mathscr{K}}(k) - Q \times \{0\}.$$

Applying Theorem 11.2 we deduce that for some $\lambda \geqq 0$, $\mu \in N_D(L(\bar{y}))$, $\bar{p} \in N_K(A_1(\bar{y}))$ not all equal to zero and all $w \in B_Y$

$$\lambda J_1'(\bar{y})(w, A w + \varphi'(\bar{y}) w) + \langle L'(\bar{y})^* \mu, w \rangle + \langle A^* \bar{p} + \varphi'(\bar{y})^* \bar{p}, w \rangle \geqq 0.$$

Hence

$$A^*\left(\lambda \frac{\partial J}{\partial u}(\bar{y}, \bar{u}) + \bar{p}\right) + \varphi'(\bar{y})^* \left(\lambda \frac{\partial J}{\partial u}(\bar{y}, \bar{u}) + \bar{p}\right) + \lambda \frac{\partial J}{\partial y}(\bar{y}, \bar{u}) + L'(\bar{y})^* \mu = 0.$$

Setting $p = -\lambda \frac{\partial J}{\partial u}(\bar{y}, \bar{u}) - \bar{p}$ we obtain (70), (71). From (70) follows that $A^* p \in C_0(\Omega)^*$ and, consequently, that for all $0 < s < \frac{n}{n-1}$, $p \in W_0^{1, s}(\Omega)$.

12. SMALL TIME LOCAL CONTROLLABILITY

Let U be a complete metric space and $f: \mathbf{R}^n \times U \to \mathbf{R}^n$ be a continuous function, $x_0 \in \mathbf{R}^n$.

We assume here that:

(*a*) f is locally Lipschitz in the first variable, uniformly on U (*see* Section 10);

(*b*) For some $\bar{u} \in U$, $f(x_0, \bar{u}) = 0$, *i.e.* x_0 is an equilibrium;

(*c*) For all x near x_0, $f(x, U)$ is a convex, compact subset of $\mathbf{R}^n$.

Consider the control system (43). It is called small time locally controllable (s. t. l. c.) if

$$\forall \mathrm{T}>0, \quad x_0 \in \operatorname{Int} \mathrm{R}(\mathrm{T}) \tag{73}$$

where R (T) is the reachable set of (43) at time $\mathrm{T} \geqq 0$.

DEFINITION 12.1. – *A vector* $v \in \mathbf{R}^n$ *is called a variation of* R(.) (*of order* $k \geqq 1$) *if for all* $t \geqq 0$

$$x_0 + t^k v \in \mathrm{R}(t) + o(t^k).$$

In other words v is a variation of order k if and only if

$$\lim_{t \to 0+} \operatorname{dist}\left(v, \frac{\mathrm{R}(t) - x_0}{t^k}\right) = 0$$

or equivalently if there exists a selection $r(t) \in \mathrm{R}(t)$ (*in general discontinuous*) *such that*

$$\frac{r(t) - x_0}{t^k} \to v.$$

THEOREM 12.2. – *Assume that for some variations* $v_1, \ldots, v_p$ *of order* $\leqq k$

$$0 \in \operatorname{Int} co\{v_1, \ldots, v_p\}. \tag{74}$$

Then (43) *is s. t. l. c. and there exist* $\mathrm{L}>0$, $\varepsilon>0$ *such that for all small* $t>0$, *all* y_1 *near* x_0 *and all* $y \in \mathrm{R}(t)$ *there exists* t_1 *such that*

$$y_1 \in \mathrm{R}(t_1), \qquad |t_1 - t| \leqq \mathrm{L} \|y_1 - y\|^{1/k}$$

Proof. – Define the set-valued map $\mathrm{G}: \mathbf{R} \to \mathbf{R}^n$ by

$$\mathrm{G}(t) = \begin{cases} \mathrm{R}(t), & t \geqq 0 \\ \varnothing, & \text{otherwise.} \end{cases}$$

There exists $\mathrm{T}>0$ such that the map G restricted to [0, T] has a closed graph (*see* for example [2]). It was shown in [19] that $v_1, \ldots, v_p \in \mathrm{G}^k(0, x_0)$. Theorem 7.2 ends the proof. □

In [19] and [23] we illustrated how the above theorem can be applied to study small time local controllability of the implicit dynamical system (39) and the differential inclusion (40).

13. APPLICATIONS TO NONSMOOTH ANALYSIS

The main aim of this section is to show how stability of the uniform open mapping principle can be exploited when one deals with nonsmooth problems. Consider a locally Lipschitz map $f: \mathbf{R}^n \to \mathbf{R}^m$. Its generalized Jacobian at a point $x_0 \in \mathbf{R}^n$ is defined by

$$\partial f(x_0) = \overline{co}\left\{ \lim_{x \to x_0} f'(x) \right\}$$

where limits are taken over all sequences $\{x_i\}$ converging to x_0 such that the derivative $f'(x_i)$ does exist and $\{f'(x_i)\}$ is a converging sequence [9]. In [42] the following approximation of f was considered: Let $\psi: \mathbf{R}^n \to [0, 1]$ be a C^∞-function having its support in the unit ball. For all integer $i \geqq 1$ define $f_i: \mathbf{R}^n \to \mathbf{R}^m$ by

$$f_i(x) = \int_{\mathbf{R}^n} f\left(x - \frac{1}{i} y\right) \psi(y)\, dy.$$

Then $f_i \in C^\infty$, $f_i \to f$ uniformly on compact sets and

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \qquad I \geqq 1$$

such that

$$\forall i \geqq I, \quad \forall x \in B_\delta(x_0), \qquad f'_i(x) \in \partial f(x_0) + \varepsilon B$$

(*see* [42]).

The result below extends the inverse function theorem from [9], p. 253.

THEOREM 13.1. – *Assume that every* $A \in \partial f(x_0)$ *is surjective, then for some* $L > 0$ *and for all* (x, y) *near* $(x_0, f(x_0))$

$$\operatorname{dist}(x, f^{-1}(y)) \leqq L \, \| f(x) - y \|.$$

Remark. – If $\rho > 0$ is such that for every $A \in \partial f(x_0)$, $\rho B \subset A(B)$ then the constant L in the above theorem can be taken equal to $\frac{2}{\rho}$. This is an easy consequence of the proof given below. □

Proof. – Pick $A \in \partial f(x_0)$. It is surjective and therefore for some $\rho' > 0$, $\rho' B \subset A(B)$. Then there exists a neighborhood $\mathcal{N}$ of A such that for every $A' \in \mathcal{N}$, $\frac{\rho'}{2} B \subset A'(B)$. Since $\partial f(x_0)$ is compact for some $\rho > 0$, $\varepsilon > 0$ and all $A \in \partial f(x_0) + \varepsilon B$, we have $\rho B \subset A(B)$. Consider f_i defined as above and let $\delta > 0$ be such that for all large i

$$\forall x \in B_\delta(x_0), \quad f'_i(x) \subset \partial f(x_0) + \varepsilon B.$$

Since $f'_i(x) B \subset f_i^{(1)}(x, f_i(x))$, for all large i and all $x \in B_\delta(x_0)$ we have $\rho B \subset f_i^{(1)}(x, f_i(x))$. On the other hand, using that f_i are equilipschitzian on $B_\delta(x_0)$, we prove that for some $0 < \bar{\delta} < \delta$ and for all $x \in B_{\bar{\delta}}(x_0)$, $\| f_i(x) - f_i(x_0) \| \leqq \frac{\delta}{4}$. From Theorem 6.1 we deduce that for some $\varepsilon > 0$

and for all large i

$$\forall x \in \mathrm{B}_\varepsilon(x_0), \quad h \in [0, \varepsilon], \qquad f_i(x) + ph\,\mathring{\mathrm{B}} \subset f_i(\mathrm{B}_h(x)).$$

This and Theorem 4.1 imply that the above holds true as well with f_i replaced by f for all x near x_0 and for all small $h>0$. Theorem 2.2 ends the proof. □

Remark. – In this paper we did not address the question of univocity of the inverse. When $m=n$ then, using the mean value theorem, it is possible to check that f^{-1} is actually single valued on a neighborhood of x_0. (see [9]). □

THEOREM 13.2 (Implicit function theorem). – *Consider a function* $g : \mathbf{R}^n \times \mathbf{R}^p \to \mathbf{R}^k$ *and let* $(\bar{x}, \bar{p}) \in \mathbf{R}^n \times \mathbf{R}^p$ *be such that* $g(\bar{x}, \bar{p}) = 0$. *Let* π *denote the projection of* $\mathbf{R}^n \times \mathbf{R}^p \times \mathbf{R}^k$ *on* $\mathbf{R}^n \times \mathbf{R}^k$. *Assume that every* $\mathrm{A} \in \pi\,\partial g(\bar{x}, \bar{p})$ *is surjective. Then there exists* $\mathrm{L}>0$ *such that for every* (x, p) *near* $(\bar{x}, \bar{p})$ *satisfying* $g(x, p)=0$ *and every* p' *near* $\bar{p}$ *there exists* x' *with*

$$g(x', p') = 0, \qquad \|x' - x\| \leqq \mathrm{L}\,\|p' - p\|.$$

To prove the above it is enough to apply Theorem 13.1 to the function $f(x, p) = (p, g(x, p))$.

Remark. – The last theorem allows to obtain results analogous to those of Section 9, concerning Lipschitz realisation of the implicit dynamical system (39) with the function f merely locally Lipschitz. □

We study next a nonsmooth control problem:

Consider a separable metric space U and a continuous function $f : \mathbf{R}^n \times \mathrm{U} \to \mathbf{R}^n$. We associate with it the control system

$$\left.\begin{aligned} x'(t) = f(x(t), u(t)), \qquad & u \in \mathscr{U} \\ x(0) = x_0 \qquad & \end{aligned}\right\} \tag{75}$$

where $\mathscr{U}$ denotes the set of all measurable selections $u : [0, 1] \to \mathrm{U}$. We consider $\mathscr{U}$ with the metric from Section 10.

Let $g : \mathbf{R}^n \to \mathbf{R}^k$ be a locally Lipschitz function and K_0, $\mathrm{K}_1 \subset \mathbf{R}^n$ be given closed sets. We study properties of the above control system under additional constraints

$$x(0) \in \mathrm{K}_0, \qquad x(1) \in \mathrm{K}_1 \tag{76}$$

and the following assumptions on datas:

(*a*) $\forall x \in \mathbf{R}^n$, $f(x, \mathrm{U})$ is bounded;

(*b*) $\forall \bar{x} \in \mathbf{R}^n$ there exists a neighborhood $\mathscr{N}$ and $\mathrm{L}>0$ such that for every $u \in \mathrm{U}$, $f(., u)$ is L-Lipschitz on $\mathscr{N}$.

THEOREM 13.3. – *Let* $(z, \bar{u})$ *be a trajectory-control pair of* (75) *defined on the time interval* [0, 1] *and satisfying the end point constraints* $z(0) \in K_0$,

$z(1) \in K_1$. *Then at least one of the following two statements holds true:*

(i) *There exist* $\lambda \in \mathbf{R}^k$ *and an absolutely continuous function* $p: [0, 1] \to \mathbf{R}^n$ *not both equal to zero, satisfying the adjoint inclusion*

$$-p'(t) \in \partial_x f(z(t), \bar{u}(t))^* p(t) \quad \text{a.e. } in\ [0, 1] \tag{77}$$

where $\partial_x f(x, u)$ *denotes the generalized Jacobian of* $f(., u)$ *at* x, *the maximum principle*

$$\max_{u \in U} \langle p(t), f(z(t), u) \rangle = \langle p(t), f(z(t), \bar{u}(t)) \rangle \quad \text{a.e. } in\ [0, 1] \tag{78}$$

and the transversality condition

$$p(0) \in N_{K_0}(z(0)), \ -p(1) \in \partial g(z(1))^* \lambda + N_{K_1}(z(1)). \tag{79}$$

(ii) *There exist* $L > 0$, $\varepsilon > 0$ *such that for all* $(a, b, c, x_0, u) \in \mathbf{R}^k \times \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n \times \mathcal{U}$ *satisfying*

$$\| a - g(z(1)) \| \leqq \varepsilon, \quad \| b \| \leqq \varepsilon, \quad \| c \| \leqq \varepsilon, \quad \| z(0) - x_0 \| \leqq \varepsilon,$$
$$\mu(\{ t \mid u(t) \neq \bar{u}(t) \}) \leqq \varepsilon$$

there exists a trajectory-control pair $(x_{\hat{u}}, \hat{u}) \in W^{1,1}(0, 1) \times \mathcal{U}$ *of* (75) *with*

$$g(x_{\hat{u}}(1)) = a, \quad x_{\hat{u}}(0) \in b + K_0, \quad x_{\hat{u}}(1) \in c + K_1$$

and if x_u *denotes the trajectory of* (75) *then*

$$\| x_{\hat{u}}(0) - x_0 \| + \mu(\{ t \mid u(t) \neq \hat{u}(t) \}) \leqq L\, (\| a - g\, (x_u(1)) \|$$
$$+ \operatorname{dist}(x_0 - b, K_0) + \operatorname{dist}(x_u(1) - c, K_1))$$

In particular this yields that for every $a \in \mathbf{R}^k$ *with* $\| a - g(z(1)) \| \leqq \varepsilon$ *and every trajectory-control pair* (x_u, u) *of* (75) *satisfying the end point constraints* (76) *with* $\| x_u(0) - z(0) \| + \mu(\{ t \in [0, 1] \mid u(t) \neq \bar{u}(t) \}) \leqq \varepsilon$ *there exists a control* $\hat{u}$ *and an initial condition* $\hat{x}_0 \in K_0$ *such that*

$$\| \hat{x}_0 - x_u(0) \| + \mu(\{ t \in [0, 1] \mid \hat{u}(t) \neq u(t) \}) \leqq L \| a - g(x_u(1)) \|$$

and the corresponding trajectory $x_{\hat{u}}(1) \in K_1$, $g(x_{\hat{u}}(1)) = a$. *Consequently if* $g(z(1))$ *is a boundary point of the set*

$$\{ g(x(1)) \mid x \text{ is a trajectory of } (75), (76) \}$$

then the statement (i) *holds true.*

A statement concerning boundary points was proved in [9], p. 200 under somewhat different constraints. The above result shows that the maximum principle (77)-(79) is verified for every trajectory-control pair $(z, \bar{u})$ where the system is not controllable in the sense of (ii).

Proof. — Denote by $x(.; u, x_0)$ the solution of (75) corresponding to the control u and the initial condition x_0. We define the single-valued map $\Phi: \mathbf{R}^n \times \mathcal{U} \to \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$ and the set-valued map G:

$\mathbf{R}^n \times \mathcal{U} \to \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$ by

$$\forall x_0 \in \mathbf{R}^n, \quad u \in \mathcal{U}, \qquad \Phi(x_0, u) = (g(x(1; u, x_0)), x_0, x(1; u, x_0)),$$
$$G(x_0, u) = \Phi(x_0, u) - \{0\} \times K_0 \times K_1.$$

By Gronwall's inequality, Φ is locally Lipschitz and therefore Graph (G) is closed. We shall use the result from Section 4. Let $\psi: \mathbf{R}^n \to [0, 1]$ be a C^∞-function having its support in the unit ball. Define $f_i: \mathbf{R}^n \times U \to \mathbf{R}^n$, $g_i: \mathbf{R}^n \to \mathbf{R}$ by

$$f_i(x, u) = \int f\left(x - \frac{1}{i} y, u\right) \psi(y)\, dy, \qquad g_i(x) = \int g\left(x - \frac{1}{i} y\right) \psi(y)\, dy$$

and denote by $x_i(.; u, x_0)$ the solution of (75) with f replaced by f_i corresponding to the control $u(.)$. Define the single-valued maps $\Phi_i: \mathbf{R}^n \times \mathcal{U} \to \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$ and the set-valued maps $G_i: \mathbf{R}^n \times \mathcal{U} \to \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$ by

$$\Phi_i(x_0, u) = (g_i(x_i(1; u, x_0)), x_0, x_i(1; u, x_0)),$$
$$G_i(x_0, u) = \Phi_i(x_0, u) - \{0\} \times K_0 \times K_1.$$

For the same reasons as before Graph (G_i) are closed sets. Since for every $u \in U$, $\{f_i(., u)\}$, $\{g_i\}$ converge uniformly on compact sets to $f(., u)$ and g respectively, we deduce from Gronwall's inequality that for every $\delta > 0$ and all $\lambda > 0$, there exists $I \geqq 1$ such that

$$\forall i \geqq I, \quad \forall x_0 \in B_\delta(z(0)), \quad \forall u \in B_\delta(\bar{u}), \qquad G_i(x_0, u) \subset G(x_0, u) + \lambda B. \quad (80)$$

If there exist $\varepsilon > 0$, $I \geqq 1$ such that for some $\rho > 0$, $M > 0$ and all $i \geqq I$ we have

$$\rho B \subset \bigcap_{\substack{(c, d) \in \mathrm{Graph}\,(G_i) \\ (c, d) \in B_\varepsilon(z(0), \bar{u}) \times B_\varepsilon(g(z(1)), 0, 0)}} \overline{co}\,(G_i^{(1)}(c, d) \cap MB). \quad (81)$$

Then, by Theorem 6.5 and (81) for some $\delta > 0$ and for all large i

$$\forall (c, d) \in \mathrm{Graph}\,(G_i) \cap B_\delta(z(0), \bar{u}) \times B_\delta(g(z(1)), 0, 0),$$
$$h \in [0, \delta], \qquad d + \rho h \mathring{B} \subset G_i(B_h(c))$$

and by Theorem 4.1 the above holds true with G_i replaced by G and δ replaced by $\frac{\delta}{4}$. Hence from Theorem 2.1 and definition of G we deduce that statement (ii) is verified.

Let us assume next that there exists no ε, ρ, $M > 0$ satisfying (81). Then, by the separation theorem, we can find $y_i \to z(0)$, $u_i \to \bar{u}$, $a_i \to_{K_0} z(0)$, $b_i \to_{K_1} z(1)$, $j_i \to \infty$, $N_i \to \infty$ such that for some $(\lambda_i, \eta_i, q_i) \in \mathbf{R}^k \times \mathbf{R}^n \times \mathbf{R}^n$ of $\|(\lambda_i, \eta_i, q_i)\| = 1$ and $\mu_i \to 0+$ we have

$$\left.\begin{aligned} &\forall (\alpha, \beta, \gamma) \in \overline{co}\,(G_{j_i}^{(1)}(y_i, u_i, \Phi_{j_i}(y_i, u_i) - (0, a_i, b_i)) \cap N_i B), \\ &\langle \lambda_i, \alpha \rangle + \langle \eta_i, \beta \rangle + \langle q_i, \gamma \rangle \geqq -\mu_i. \end{aligned}\right\} \quad (82)$$

From now on the proof reminds in many aspects those from [41] and [22]. So we shall omit many details. Let $S_i(.; s)$ denote the fundamental solution of the linear system

$$Z' = \frac{\partial f_{j_i}}{\partial x}(x_{j_i}(t; u_i, y_i), u_i(t))Z; \qquad Z(s) = \mathrm{Id}.$$

Define absolutely continuous functions $p_i : [0, 1] \to \mathbf{R}^n$ by

$$p_i(t) = S_i(1; t)^* (-g'_{j_i}(x_{j_i}(1; u_i, y_i))^* \lambda_i - q_i) \tag{83}$$

Then p_i is the solution of the adjoint system

$$\begin{aligned} -p'_i &= \frac{\partial f_{j_i}}{\partial x}(x_{j_i}(t; u_i, y_i), u_i(t))^* p_i; \\ -p_i(1) &= g'_{j_i}(x_{j_i}(1; u_i, y_i))^* \lambda_i + q_i. \end{aligned} \tag{84}$$

Taking subsequences and keeping the same notations, we may assume that

$$\lim_{i\to\infty} \lambda_i = \lambda, \qquad \lim_{i\to\infty} \eta_i = \eta, \qquad \lim_{i\to\infty} q_i = q,$$
$$\lim_{i\to\infty} g'_{j_i}(x_{j_i}(1; u_i, y_i))^* = \omega$$

and that $\{p_i\}$ converges weakly in $W^{1,\infty}$ to some p. Observe that from the Gronwall inequality and the assumptions (*a*), (*b*) follows that

$$x_{j_i}(., u_i, y_i) \to z(.) \text{ uniformly on } [0, 1]. \tag{85}$$

Thus from the definition of g_i we get $\omega \in \partial g(z(1))^*$. It is clear that for all i

$$-\{0\} \times T_{K_0}(a_i) \times \{0\} \subset G_{j_i}^{(1)}(y_i, u_i, \Phi_{j_i}(y_i, u_i) - (0, a_i, b_i))$$
$$-\{0\} \times \{0\} \times T_{K_1}(b_i) \subset G_{j_i}^{(1)}(y_i, u_i, \Phi_{j_i}(y_i, u_i) - (0, a_i, b_i)).$$

Hence from (82) we deduce that

$$\forall (k, m) \in T_{K_0}(a_i) \times T_{K_1}(b_i) \quad \text{of } \|(k, m)\| \leq N_i,$$
$$\langle \eta_i, -k \rangle + \langle q_i, -m \rangle \geqq -2\mu_i. \tag{86}$$

On the other hand from [3], Chapter 4

$$\liminf_{x' \to_{K_0} z(0)} T_{K_0}(x') = C_{K_0}(z(0)), \qquad \liminf_{x' \to_{K_1} z(1)} T_{K_1}(x') = C_{K_1}(z(1)).$$

Thus, by (86), $\eta \in N_{K_0}(z(0))$ and $q \in N_{K_1}(z(1))$. From (83) follows that $-p(1) = \omega\lambda + q \in \partial g(z(1))^* \lambda + N_{K_1}(z(1))$.

Differentiating with respect to the initial condition we also prove that for every $w_0 \in \mathbf{R}^n$ of $\| w_0 \| \leqq 1$, the solution w of the system

$$w' = \frac{\partial f_{j_i}}{\partial x} (x_{j_i}(t; u_i, y_i), u_i(t)), \qquad t \in [0, 1]$$

$$w(0) = w_0$$

verifies

$$(g'_{j_i}(x_{j_i}(1; u_i, y_i))\, w(1), w_0, w(1)) \in G_{j_i}^{(1)}(y_i, u_i, \Phi_{j_i}(y_i, u_i) - (0, a_i, b_i)).$$

Since it is given by $w(t) = S_i(t; 0)\, w_0$, we deduce from (82) that for every $w_0 \in \mathbf{R}^n$ of $\| w_0 \| \leqq 1$ and all large i

$$\langle g'_{j_i}(x_{j_i}(1; u_i, y_i))^* \lambda_i, S_i(1; 0)\, w_0 \rangle + \langle \eta_i, w_0 \rangle + \langle q_i, S_i(1; 0)\, w_0 \rangle = \langle -p_i(0) + \eta_i, w_0 \rangle \geqq -\mu_i.$$

Taking the limit we obtain that for every $w_0 \in \mathbf{R}^n$, such that $\| w_0 \| \leqq 1$, $\langle -p(0) + \eta, w_0 \rangle \geqq 0$.

Therefore $p(0) = \eta \in N_{K_0}(z(0))$ and we proved (79). Let M be the Lipschitz constant of Φ and Φ_i on a neighborhood of $(z(0), \bar{u})$ (it exists because of the Gronwall inequality and the definition of f_i, g_i). Set

$$V_i(t) = \overline{co}\, f_{j_i}(x_{j_i}(t; u_i, y_i), U) - f_{j_i}(x_{j_i}(t; u_i, y_i), u_i(t)).$$

From Section 10.1 we deduce that every solution w of the linear control system

$$w' = \frac{\partial f_{j_i}}{\partial x} (x_{j_i}(t; u_i, y_i), u_i(t))\, w + y(t), \; y(t) \in V_i(t)$$

$$w(0) = 0$$

verifies

$$(g'_{j_i}(x_{j_i}(1; u_i, y_i))\, w(1), 0, w(1)) \in \overline{co}\,(G_{j_i}^{(1)}(y_i, u_i, \Phi(y_i, u_i) - (0, a_i, b_i)) \cap MB)$$

and from (82) we deduce that for all large i and for every measurable selection $v(t) \in V_i(t)$

$$\left\langle g'_{j_i}(x_{j_i}(1; u_i, y_i))^* \lambda_i + q_i, \int_0^1 S_i(1; t)\, v(t)\, dt \right\rangle = \int_0^1 \langle -p_i(t), v(t) \rangle\, dt \geqq -\mu_i$$

Since $f_i(., u) \to f(., u)$ uniformly on compact sets, using (85), we obtain from the last inequality that for every measurable selection $v(t) \in \overline{co}\, f(z(t), U) - f(z(t), u(t))$, $\int_0^1 \langle -p(t), v(t) \rangle\, dt \geqq 0$. Hence p satisfies the maximum principle (78). Finally from (85), (84) and Mazur's lemma we deduce that p is a solution of the adjoint equation (77). □

REFERENCES

[1] J.-P. AUBIN, Lipschitz Behavior of Solutions to Convex Minimization Problems, *Math. Operat. Res.*, Vol. **9**, 1984, pp. 87-111.

[2] J.-P. AUBIN and A. CELLINA, *Differential Inclusions*, Springer-Verlag (Grundlehren der Math. Wissenscharften), Vol. **264**, 1984.

[3] J.-P. AUBIN and H. FRANKOWSKA, *Set-Valued Analysis*, Birkhäuser, Berlin, Boston, 1990.

[4] J.-P. AUBIN and H. FRANKOWSKA, On inverse function theorems for set-valued maps, *J. Math. Pure Appl.*, Vol. **66**, 1987, pp. 71-89.

[5] P. BEAUZAMY, *Introduction to Banach spaces and their Geometry*, North Holland Mathematics study, Vol. 68, 1985.

[6] R. M. BIANCHINI and G. STEFANI, *Sufficient Conditions of Local Controllability*, Preprint Istituto Matematico Ulisse Dini 1986.

[7] J. F. BONNANS and E. CASAS, Optimal Control of Semilinear Multistate Systems with State Constraints, *SIAM J, Control Optimi.*, Vol. 27, 1989, pp. 446-455.

[8] F. CLARKE, On the Inverse Function Theorem, *Pac. J. Math.*, Vol. **64**, 1976, pp. 97-102.

[9] F. CLARKE, *Optimization and Nonsmooth Analysis*, Wiley Interscience, 1983.

[10] B. D. CRAVEN and M. Z. NASHED, Generalized Implicit Function Theorems when the Derivative has no Bounded Inverse, *Nonlinear Analysis*, Vol. **6**, 1982, pp. 375-387.

[11] A. V. DMITRUK, A. A. MILYUTIN and N. P. OSMOLOVSKII, Ljusternik's Theorem and the Theory of Extrema, *Uspekhi Mat. Nauk.*, Vol. **35:6**, 1980, pp. 11-46; *Russian Math. Surveys*, Vol. **35:6**, 1980, pp. 11-51.

[12] I. EKELAND, On the Variational Principle, *J. Math. Anal. Appl.*, Vol. **47**, 1974, pp. 324-358.

[13] I. EKELAND, Nonconvex Minimization Problems, *Bull. Am. Math. Soc.*, Vol. **1**, 1979, pp. 443-474.

[14] H. FATTORINI, A Unified Theory of Necessary Conditions for Nonlinear Non-convex Control Systems, *Applied Math. Optimiz.*, Vol. **15**, 1986, pp. 141-185.

[15] H. FATTORINI and H. FRANKOWSKA, Necessary Conditions for Infinite Dimensional Control Problems, *Mathematics of Systems, Signals and Control*, 1991.

[16] H. FRANKOWSKA, Théorèmes d'application ouverte et de fonction inverse, *C. R. Acad. Sci., Paris*, Vol. **305**, 1987, pp. 773-776.

[17] H. FRANKOWSKA, High Order Inverse Function Theorems, Analyse non linéaire, *Proceedings of the Conference on Applied Nonlinear Analysis*, Perpignan, June 22-26, 1987, Gauthier-Villars, Paris, 1989.

[18] H. FRANKOWSKA, An Open Mapping Principle for Set-valued Maps, *J. Math. Anal. Appl.*, Vol. **127**, 1987, pp. 172-180.

[19] H. FRANKOWSKA, Local Controllability of Control Systems with Feedback, *J. Optimiz. Theor. Appl.*, Vol. **60**, 1989, pp. 277-296.

[20] H. FRANKOWSKA, Local Controllability and Infinitesimal Generators of Semi-groups of Set-valued Maps, *SIAM J., Control Optimiz.*, Vol. **25**, 1987, pp. 412-432.

[21] H. FRANKOWSKA, The Maximum Principle for an Optimal Solution to a Differential Inclusion with end Point Constraints, *SIAM J., Control Optimiz.*, Vol. **25**, 1987, pp. 145-157.

[22] H. FRANKOWSKA, The First Order Necessary Conditions for Nonsmooth Variational and Control Problems, *SIAM J., Control Optimiz.*, Vol. **22**, 1984, pp. 1-12.

[23] H. FRANKOWSKA, Set-valued Analysis and Some Control Problems, Modern Optimal Control, *Proceedings of the International Conference "30 years of Modern Control Theory"*, Kingston, June 3-6, 1988, E. Roxin Editor, Marcel Dekker, New York, Basel, 1989.

[24] L. M. GRAVES, Some Mapping Theorem, *Duke Math. J.*, Vol. **17**, 1950, pp. 111-114.
[25] K. GRASSE, A Higher-order Sufficient Condition for Local Surjectivity, *Nonlinear Analysis*, Vol. **10**, 1986, pp. 87-96.
[26] A. D. IOFFE, Regular Points of Lipschitz Functions, *Trans. Am. Math. Soc.*, Vol. **251**, 1979, pp. 61-69.
[27] A. D. IOFFE, Nonsmooth Analysis: Differential Calculus of Nondifferentiable Mappings, *Trans. Am. Math. Soc.*, **266**, 1981, pp. 1-56.
[28] A. E. IOFFE, On the Local Surjection property, *Nonlinear Analysis*, Vol. **11**, 1987, pp. 565-592.
[29] C. KURATOWSKI, *Topologie*, Panstwowe Wydawnictwo Naukowe, 1958.
[30] E. B. LEE and L. MARKUS, *Foundations of Optimal Control Theory*. John Wiley, New York 1967.
[31] L. A. LJUSTERNIK, Conditional Extrema of Functionals, *Mat. Sb.*, Vol. **41**, 1934, pp. 390-401.
[32] L. A. LJUSTERNIK and V. I. SOBOLEV, *Elementy Funktsionalnogo Analiza*, Second ed., Nauka Moskow, 1965 (in Russian).
[33] K. MAGNUSSON, A. J. PRITCHARD and M. D. QUINN, *The Application of Fixed Point to Global Nonlinear Controllability Problems*, Mathematical Control Theory, Banach Center Publications, Vol. **14**, 1985, pp. 319-344.
[34] K. MAGNUSSON and A. J. PRITCHARD, *Local Exact Controllability of Non-linear Evolution Equations*, Recent Advances in Differential Equations, Academic Press, 1981, Vol. **271**, pp. 271-280.
[35] J. NECAS, *Les méthodes directes en théorie des équations elliptiques*, Masson, Paris, 1967.
[36] A. PAZY, *Semi-groups of linear operators and applications to partial differential equations*, Spinger, Berlin, 1978.
[37] S. ROBINSON, Perturbed Kuhn-Tucker Points and Rates of Convergence for a Class of Nonlinear-programming Algorithms, *Math. Programming*, Vol. **7**, 1974, pp. 1-16.
[38] R. T. ROCKAFELLAR, Lipschitzian Properties of Multifunctions, *Non-linear Analysis*, T.M.A., 9, 1985, pp. 867-885.
[39] H. SUSSMANN, A Sufficient Condition for Local Controllability, *SIAM J., Control Optimiz.*, Vol. **16**, 1978, pp. 790-802.
[40] J. TREIMAN, Characterization of Clarke's Tangent and Normal Cones in Finite and Infinite Dimension, *Nonlinear Analysis*, Vol. **7**, 1983, pp. 771-783.
[41] J. WARGA, Necessary Conditions Without Differentiability Assumptions in Optimal Control, *J. Differ. Equations*, Vol. **18**, 1975, pp. 41-62.
[42] J. WARGA, *Derivate Containers, Inverse Functions and Controllability*, in *Calculus of Variations and Control Theory*, D. L. RUSSEL, Ed., Academic Press, New York, 1976, pp. 13-46.
[43] J. WARGA, Optimization and Controllability Without Differentiability Assumptions, *SIAM J. Control Optimiz.*, Vol. **21**, 1983, pp. 837-855.

(*Manuscript received March* 23, 1989;
in revised form May 11, 1989.)

La Revue *Annales de l'Institut Henri Poincaré/Analyse non linéaire* est une publication de Gauthier-Villars, société Anonyme, constituée pour 99 ans, au capital de 3 089 600 F.

Siège social : 17, rue Rémy-Dumoncel, Paris, P.D.G. : J. LISSARRAGUE.

Actionnaire : Bordas S. A. (99,8 % des parts).

Directeur de la publication : J. LISSARRAGUE. Responsable de la Rédaction : I. EKELAND.