LIPSCHITZ BEHAVIOR OF SOLUTIONS TO CONVEX MINIMIZATION PROBLEMS

Jean-Pierre Aubin

June 1981
WP-81-76

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

ABSTRACT

We derive the Lipschitz dependence of the set of solutions of a convex minimization problem and its Lagrange multipliers upon the natural parameters from an Inverse Function Theorem for set-valued maps. This requires the use of contingent and Clarke derivatives of set-valued maps, as well as generalized second derivatives of convex functions.

LIPSCHITZ BEHAVIOR OF SOLUTIONS
TO CONVEX MINIMIZATION PROBLEMS

Jean-Pierre Aubin

INTRODUCTION

1. One of the objectives of this paper is to propose a solution to the much studied problem of the local Lipschitz dependence of the set of solutions of a (convex) minimization problem and its Lagrange multipliers upon the natural parameters of this problem.

Namely, consider two finite dimensional vector spaces X and Y, a linear operator A_0 from X onto Y and two proper lower semi-continuous convex functions $U : X \to \mathbb{R} \cup \{+\infty\}$ and $V : Y \to \mathbb{R} \cup \{+\infty\}$.

The parameters of the optimization problem are the operator A_0 and the vectors $y_0 \in Y$ and $p_0 \in X^*$ belonging to

$$(1) \qquad \begin{cases} \text{i)} & y_0 \in \text{Int}\ (\text{Dom}\ V - A_0\ \text{Dom}\ U) \\ \text{ii)} & p_0 \in \text{Int}\ (A_0^*\ \text{Dom}\ V^* + \text{Dom}\ U^*) \end{cases}$$

where $A_0^* \in L(Y^*, X^*)$ is the transpose of A_0 and U^* and V^* are the conjugate functions of U and V respectively.

Then we know that there exist solutions $x_0 \in X$ and $q_0 \in Y^*$ to the minimization problems

$$(2)\quad \begin{cases} U(x_0) + V(A_0x_0 + y_0) - \langle p_0, x_0\rangle \\ = \min\limits_{x\in X} (U(x) + V(A_0x + y_0) - \langle p_0, x\rangle) \end{cases}$$

$$(3)\quad \begin{cases} U^*(-A_0^*q_0 + p_0) + V^*(q_0) - \langle q_0, y_0\rangle \\ = \min\limits_{q\in Y^*} (U^*(-A_0^*q + p_0) + V^*(q) - \langle q, y_0\rangle) \end{cases}$$

related by the equation

$$(4)\quad \begin{cases} U(x_0) + U^*(-A_0^*q_0 + p_0) + V(A_0x_0 + y_0) + V^*(q_0) \\ = \langle p_0, x_0\rangle + \langle q_0, y_0\rangle . \end{cases}$$

We denote by $F^{-1}(p_0, y_0, A_0)$ the set of pairs (x_0, q_0) satisfying the three properties (2), (3) and (4), i.e., of pairs of solutions x_0 to the minimization problem (2) and its Lagrange multipliers q_0. We state the following problems:

a) Do there exist neighborhoods $\mathcal{U}$ of (x_0, q_0) and $\mathcal{V}$ of (p_0, y_0, A_0) such that

$\forall (p,y,A) \in \mathcal{V},\quad F^{-1}(p,y,A) \cap \mathcal{U} \neq \phi$?

b) Does the map $(p,y,A) \in \mathcal{V} \to F^{-1}(p,y,A)$ possess a Lipschitz behavior?

c) Can we find the effect of marginal variations δp, δy and δA on the solution x_0 and its Lagrange multiplier q_0 ?

For solving these problems, we shall define a suitable concept of generalized second derivative of convex functions U: for each x_0 and $p_0 \in \partial U(x_0)$, the second derivative $\partial^2 U(x_0, p_0)$ is a monotone closed convex process from X to itself, i.e., a set valued map whose graph is a monotone closed convex cone. Such maps are set valued analogs of positive continuous linear operators. Naturally, if U is twice continuously differentiable,

$\partial^2 U(x_0,p_0)$ coincides with the second derivative in the usual sense.

We shall prove the following result. If the monotone closed convex process from $X \times Y$ to itself defined by the matrix

$$\begin{pmatrix} \partial^2 U(x_0,p_0 - A_0^* q_0) & A_0^* \\ -A_0 & (\partial^2 V(Ax_0+y_0,q_0))^{-1} \end{pmatrix}$$

is surjective, then there exist neighborhoods $\mathcal{U}$ of (x_0,q_0) and $\mathcal{V}$ of (p_0,y_0,A_0) such that, for all parameters (p,y,A) in $\mathcal{V}$, the set $F^{-1}(p,y,A)\cap\mathcal{U}$ of solutions (x,q) to the problems (2), (3) and (4) is nonempty. This set of solutions depends upon the parameters p, y and A in a Lipschitz way. The marginal variations of the parameters δp, δy and δA and the associated variations δx and δq are related by

$$\begin{pmatrix} \delta x \\ \delta q \end{pmatrix} \in \begin{pmatrix} \partial^2 U(x_0,p_0 - A_0^* q_0) & A_0^* \\ -A_0 & (\partial^2 V(Ax_0 + y_0,q_0))^{-1} \end{pmatrix}^{-1} \begin{pmatrix} \delta p - \delta A^* q_0 \\ \delta y + \delta A \cdot x_0 \end{pmatrix}$$

We shall arrive at this result by building a quite natural machinery and by piecing together independent results which may have intrinsic values.

2. We already observed that the solution of the third problem requires a convenient definition of a derivative of the set-valued map F, because we cannot assume uniqueness of the solutions without restricting too much the validity of the result.

We also observe that the generalized gradient of a convex function U is a set-valued map $x \to \partial U(x)$. Therefore, the definition of a generalized second derivative requires a suitable concept of a derivative of a set-valued map.

Finally, the only available strategy for solving the above problems is to define the set $F^{-1}(p_0,y_0,A_0)$ of solutions as the set of solutions to the equivalent (hamiltonian) system of inclusions

$$(5)\quad \begin{cases} \text{i)} & p_0 \in \partial U(x_0) + A_0^* q_0 \\ \text{ii)} & y_0 \in -A_0 x_0 + \partial V^*(q_0) \end{cases}$$

and use a sort of Inverse Function Theorem for the "matrix" of set-valued maps

$$\begin{pmatrix} \partial U & A_0^* \\ -A_0 & \partial V^* \end{pmatrix}$$

mapping $X \times Y^*$ to $X^* \times Y$.

This is a third reason for introducing one or several concepts of derivatives of set-valued maps which allow to state Inverse Function Theorems for set-valued maps involving reasonable and checkable assumptions.

Inverse Function Theorems for nondifferentiable maps or for special set-valued maps are already known and widely used. Let us mention, among other papers, the paper of Clarke [3] using generalized Jacobians, the papers of Ioffe [1], [2] using fans and a series of papers of Robinson [1], [2], [6], [7], [9] studying inverse functions for sums of differentiable maps and maximal monotone operators. Robinson used his results in [9] for studying the dependence of the set of solutions upon parameters and Cornet & Laroque [1] used the Clarke Inverse Function Theorem for solving the above problems for optimization problems relevant to economic theory.

3. In this paper, we use an approach due to Ekeland for proving the Inverse Function Theorem which involves his powerful theorem (see Ekeland [1]). This approach was used in Aubin [7] for devising an Inverse Function Theorem for set-valued maps, which was both simplified and dramatically improved by Lebourg [1]. We shall adopt Lebourg's approach to suit our purpose.

So, we have to tackle the issue of defining a derivative to set-valued maps. We follow a very simple strategy, which is the ancient Fermat's geometrical view, which regards the graph of the derivative at a point as the tangent to the graph of the map.

If the graph of a single-valued map is a smooth manifold, then the tangent space at a point is a vector subspace, and thus, is the graph of a linear operator.

If the graph of a set-valued map is convex, there is still no ambiguity for defining a tangent cone to the graph, which is a closed convex cone: then it is the graph of a closed convex process (see Aubin [6]). When the graph of a set-valued map is neither smooth nor convex, we have to make a choice of a tangent cone among the many suggestions proposed in the fast growing specialized literature. We shall retain only two tangent cones, the _contingent cone_, introduced by Bouligand (see for instance Aubin [7]) and the _Clarke tangent cone_ (see for instance Clarke [1], [2] and Rockafellar [4], [5], [6] among the many papers dealing with this topic.) These two cones are closely related since the Clarke tangent cone at x_0 is some kind of limit of the contingent cones at x when x converges to x_0. Therefore, properties of the Clarke tangent cone at a point x_0 yield (weaker) properties of the contingent cones at the points of a neighborhood of x_0, properties which are most of the time sufficient to suit our purposes.

The Clarke tangent cone is always a closed _convex_ cone contained in the contingent cone. Therefore, we shall define both contingent and Clarke derivatives to a set-valued map, whose graphs are the contingent cone and the Clarke tangent cones to the graph.

Maps with closed convex graph are called _closed convex processes_ (see Rockafellar [1], [2]) which are the set-valued analogs of continuous linear operators. So, Clarke derivatives are _closed convex processes_. Closed convex processes enjoy many properties of the continuous linear operators, and, specially, the Banach open mapping principle, which plays an important underlying role in the Inverse Function Theorem. Robinson [2] and Ursescu [1] proved that the inverse of a _surjective_ closed convex process is a Lipschitz set-valued map.

The Inverse Function Theorem that we shall propose has a very simple formulation.

Let F *be a set-valued map with a closed graph and let* (x_0, y_0) *belong to the graph of* F. *Assume that the Clarke derivative of* F *at* (x_0, y_0), *which is a closed convex process, is surjective. Then there exist neighborhoods* $\mathcal{U}$ *of* x_0 *and* $\mathcal{V}$ *of* y_0 *such that* $y \in \mathcal{V} \to F^{-1}(y) \cap \mathcal{U}$ *has nonempty values and is quasi-Lipschitz.*

4. The Inverse Function Theorem is certainly as useful for other applications as the classical one. We propose in this paper to use it for "computing" the Clarke tangent cone to subsets of the form $L \cap A^{-1}(M)$ where $L \subset X$ and $M \subset Y$ are closed subsets and where A is a continuously differentiable map from X to Y. We denote by $C_K(x)$ the Clarke tangent cone to K at x.

When L and M are convex and A is linear, we know that the condition

$$(6) \qquad 0 \in \text{Int } (M - A(L))$$

implies that

$$(7) \qquad C_{L \cap A^{-1}(M)}(x) = C_L(x) \cap A^{-1} C_M(Ax)$$

(See Aubin [2], [3] and [6], for instance).

For nonconvex subsets, Rockafellar [4] has proved that the condition

$$(8) \qquad 0 \in \text{Int } C_M(Ax) - \nabla A(x) C_L(x)$$

implies that

$$(9) \qquad C_{L \cap A^{-1}(M)}(x) \supset C_L(x) \cap A^{-1} C_M(Ax) \quad .$$

The Inverse Function Theorem allows to relax assumption (8) and to replace it by the weaker assumption

(10) $$C_M(Ax) - \nabla A(x)C_L(x) = Y$$

(which is equivalent to $0 \in \text{Int}\ (C_M(Ax) - \nabla A(x)C_L(x))$. This assumption does not require that the Clarke tangent cone to M has a nonempty interior. Such formulas allow to devise a satisfying calculus for Clarke derivatives.

5. We define in a first section the contingent and Clarke tangent cones and, in the second section, the contingent and Clarke derivatives of a set-valued map. We devote the third section to the Inverse Function Theorem, which we apply in the fourth for proving the formula $C_L(x_0) \cap \nabla A(x_0)^{-1}C_M(Ax_0) \subset C_{L\cap A^{-1}(M)}(x_0)$. The fifth section deals with the proof of the regularity of the solutions of an optimization problem. Since the surjectivity of the Clarke derivative plays such an important role, we adapt in the last section the Lax-Milgram Theorem to closed convex processes.

1. CONTINGENT AND CLARKE TANGENT CONES

We recall the definitions of Bouligand's contingent cone to K at x and of the Clarke tangent cone and we mention the properties we need. Let K be a nonempty subset of a Banach space X. We denote by εB and $\varepsilon \overset{\circ}{B}$ the ball (resp. open ball) of center 0 and radius $\varepsilon > 0$. We set $B_K(x_0,\varepsilon) := K \cap (x_0 + \varepsilon B)$ and the symbol $x \underset{K}{\to} x_0$ denotes the convergence of x to x_0 in K.

<u>Definition 1</u>

We say that the subset

(1) $$T_K(x) := \bigcap_{\varepsilon>0} \bigcap_{\alpha>0} \bigcup_{0<h\leq\alpha} \left(\frac{1}{h}(K-x) + \varepsilon B\right)$$

is the "<u>contingent</u> <u>cone</u>" to K at x. ∎

In other words, $v \in T_K(x)$ if and only if

(2) $$\begin{cases} \forall \varepsilon > 0,\ \forall \alpha > 0,\ \exists u \in v + \varepsilon B,\ \exists h \in]0,\alpha] \\ \text{such that } x + hu \in K \end{cases},$$

or, equivalently, $v \in T_K(x)$ if and only if there exists a sequence of strictly positive numbers h_n and of elements $u_n \in X$ satisfying

$$(3) \quad \text{i)} \lim_{n\to\infty} u_n = v, \quad \text{ii)} \lim_{n\to\infty} h_n = 0, \quad \text{iii)} \forall n \geq 0,\ x + h_n u_n \in K .$$

We characterize the contingent cone by using the distance function $d_K(\cdot)$ to K defined by $d_K(x) := \inf \{\|x-y\| \mid y \in K\}$:

$$(4) \qquad v \in T_K(x) \text{ if and only if } \liminf_{h \to 0+} \frac{d_K(x+hv)}{h} = 0 . \ \blacksquare$$

It is quite obvious that the contingent cone is a <u>closed</u> <u>cone</u>, which is trivial when x belongs to the interior of K:

$$(5) \qquad \text{When } x \in \text{Int }(K) \ , \ \text{then } T_K(x) = X \ .$$

For all $x \in X$, we have $T_X(x) = X$. We shall set $T_\phi(x) := \phi$. ■

It is convenient to recall the definition of the "limitinf" of a family of subsets $F(u)$.

<u>Definition 2</u>

Let U be a metric space, u_0 belong to U and F be a set-valued map from U to X. We set

$$(6) \qquad \liminf_{u \to u_0} F(u) := \bigcap_{\varepsilon>0} \bigcup_{\eta>0} \bigcap_{u \in B(u_0,\eta)} (F(u) + \varepsilon B) \ .$$

We observe that

$$(7) \qquad \liminf_{u \to u_0} F(u) \subset F(u_0)$$

and that F is lower semicontinuous at u_0 if and only if

$$(8) \qquad F(u_0) = \liminf_{u \to u_0} F(u) \ .$$

Let us set $d(v,K) := \inf_{w \in K} \|v-w\|$.

It is useful to notice that v belongs to $\liminf_{u \to u_0} F(u)$ if and only if

(9) $\forall \varepsilon > 0, \exists \eta > 0$ such that $\sup_{u \in B(u_0, \eta)} d(v, F(u)) \leq \varepsilon$.

<u>Definition 3</u>

We say that the subset

(10) $$C_K(x_0) := \liminf_{\substack{h \to 0+ \\ x \to_K x_0}} \frac{1}{h}(K-x) := \bigcap_{\varepsilon>0} \bigcup_{\alpha,\beta>0} \bigcap_{\substack{x \in B_K(x_0,\alpha) \\ h \in]0,\beta]}} \left(\frac{1}{h}(K-x) + \varepsilon B\right)$$

is the Clarke tangent cone to K at x_0. (See Clarke [1]) ▲

In other words, $v \in C_K(x_0)$ if and only if

(11) $$\begin{cases} \forall \varepsilon > 0, \exists \alpha > 0, \exists \beta > 0 \text{ such that } \forall x \in B_K(x_0, \alpha) , \\ \forall h \in]0,\beta], \exists u \in v + \varepsilon B \text{ satisfying } x + hu \in K , \end{cases}$$

or, equivalently, if and only if

(12) $$\begin{cases} \text{for all sequences of elements } (x_n, h_n) \in K \times \overset{\circ}{\mathbb{R}}_+ \\ \text{converging to } (x_0, 0)\text{, there exists a sequence} \\ \text{of elements } u_n \in X \text{ converging to } v \text{ such that} \\ x_n + h_n u_n \text{ belongs to } K \text{ for all } n. \end{cases}$$

It is also characterized in the following way

(13) $$v \in C_K(x_0) \text{ if and only if } \lim_{\substack{x \to_K x_0 \\ h \to 0+}} \frac{d_K(x + hv)}{h} = 0 .$$

The Clarke tangent cone is obviously contained in the contingent cone. Actually, we observe that

(14) $$C_K(x_0) \subset \liminf_{x \to_K x_0} T_K(x) \quad .$$

If not, there would exist v_0 in $C_K(x_0)$ such that v_0 does not belong to $\liminf_{x \to_K x_0} T_K(x)$. The latter property means that there exist $\varepsilon > 0$ and a sequence of elements x_n converging to x_0 such that

$$(v_0 + \varepsilon B) \cap T_K(x_n) = \phi .$$

Then there exists a sequence of elements $h_n > 0$ converging to 0 such that

(15) $$(v_0 + \varepsilon B) \cap (\frac{1}{h_n}(K - x_n)) = \phi \quad \text{for all } n \quad .$$

Since v_0 belongs to $C_K(x_0)$, there exists a sequence of elements v_n converging to v_0 such that $x_n + h_n v_n$ belongs to K. This is a contradiction of property (15). ■

A theorem due to Cornet (see Cornet [1], [2]) states that $C_K(x_0)$ is equal to $\liminf_{x \to_K x_0} T_K(x)$ when X is finite dimensional (see also Penot [1]).

<u>Theorem 1</u>. Let K be a nonempty closed subset of a finite dimensional space. Then

(16) $$C_K(x_0) = \liminf_{x \to_K x_0} T_K(x) = \liminf_{x \to_K x_0} \overline{co}\, T_K(x)$$

▲

<u>Corollary 1</u>

Assume that X is finite dimensional.

The set-valued map $x \to T_K(x)$ is lower semicontinuous at x_0 if and only if the contingent cone to K at x_0 coincides with the Clarke tangent cone to K at x_0. ▲

For the sake of completeness, we provide Cornet's proof, which is based on the following lemmas.

Lemma 1

Let $K \subset X$ be a closed subset. We denote by $\pi_K(y)$ the subset of elements $x \in K$ such that $\|x-y\| = d_K(y)$. We obtain the following inequalities

(17) $\forall y \notin K,\ \forall x \in \pi_K(y),\ \forall v \in \overline{co}\ T_K(x)$, then $\langle y-x,v\rangle \le 0$. ▲

Proof.

Let $x \in \pi_K(y)$ and $v \in T_K(x)$. We deduce from the inequalities $\|y-x\| - d_K(x+hv) = d_K(y) - d_K(x+hv) \le \|y-x-hv\|$ that

$$\frac{\langle y-x,v\rangle}{\|y-x\|} = \lim_{h\to 0+} \frac{\|y-x\| - \|y-x-hv\|}{h} \le \liminf_{h\to 0+} \frac{d_K(x+hv)}{h} = 0$$

for $y \neq x$, $u \to \|u\|$ is differentiable at $u \neq 0$ and $v \in T_K(x)$. So $\langle y-x,v\rangle \le 0$ for all $v \in T_K(x)$, and, consequently, for all $v \in \overline{co}\ T_K(x)$.

■

Lemma 2

For any $y \in X$, we have

(18) $$\liminf_{h\to 0+} \frac{1}{2h}\left(d_K(y+hv)^2 - d_K(y)^2\right) \le d_K(y)\,d(v,\overline{co}\ T_K(\pi_K(y))) .$$ ▲

Proof.

Let us take x in $\pi_K(y)$. We observe that

$$\frac{1}{2h}\left(d_K(y+hv)^2 - d_K(y)^2\right) \le \frac{1}{2h}\left(\|y+hv-x\|^2 - \|y-x\|^2\right)$$

because $d_K(y) = \|y-x\|$. Therefore

$$\liminf_{h\to 0+} \frac{1}{2h}\left(d_K(y+hv)^2 - d_K(y)\right)^2 \le \langle y-x,v\rangle$$

and, for all $w \in \overline{co}\ T_K(x)$, we deduce from the above lemma that

$$\liminf_{h\to 0+} \frac{1}{2h}\,(d_K(y+hv)^2 - d_K(y)^2) \le \langle y-x, v-w\rangle$$

$$\le \|y-x\|\,\|v-w\| = d_K(y)\,\|v-w\| \quad .$$

Lemma 1 ensues by taking the infimum when w ranges over $\overline{co}\ T_K(x)$ and x over $\pi_K(y)$. ■

Corollary 2

Let us consider the Lipschitz function f defined by $f(t) := \frac{1}{2}d_K(x+tv)^2$. For almost all $t \ge 0$, we have

$$(19) \qquad f'(t) \le d_K(x+tv)\,d(v, \overline{co}\ T_K(\pi_K(x+tv))) \quad .$$

▲

Proof of Theorem 1

Let v_0 belong to $\liminf_{x\to x_0} \overline{co}\ T_K(x)$. Then, for all $\varepsilon > 0$, there exists $\eta > 0$ such that, for all $x \in B_K(x_0,\eta)$, $v_0 \in \overline{co}\ T_K(x) + \varepsilon B$. Now, if x belongs to $B_K(x_0,\alpha)$ and $t \in]0,\beta[$, then $\pi_K(x+tv_0) \subset B_K(x_0,\eta)$ whenever $2\alpha + \beta\|v_0\| \le \eta$. This happens for instance when $\alpha := \eta/4$ and $\beta := \eta/2\|v_0\|$. By setting $f(t) := \frac{1}{2}\,d_K(x+tv_0)^2$, we deduce from Corollary 2 that

$$f'(t) \le d_K(x+tv_0)\,d(v_0, \overline{co}\ T_K(\pi_K(x+tv_0)))$$

$$\le \varepsilon d_K(x+tv_0) \le \varepsilon t\|v_0\|$$

because $d_K(x+tv_0) \le t\|v_0\|$.

Therefore, for all $x \in B_K(x_0,\alpha)$ and $t \in]0,\beta]$,

$$\frac{1}{2}\,d_K(x+hv_0)^2 = f(h) - f(0) = \int_0^h f'(t)\,dt \le \varepsilon\|v_0\|\frac{h^2}{2}$$

and consequently,

$$\lim_{\substack{x \to_K x_0 \\ h\to 0+}} \frac{d_K(x+hv_0)}{h} = 0 \quad .$$

This implies that v_0 belongs to the Clarke tangent cone $C_K(x_0)$.

Then, by formula (14), we obtain:

$$\liminf_{x \to_K x_0} T_K(x) \subset \liminf_{x \to_K x_0} \overline{co}\, T_K(x) \subset C_K(x_0) \subset \liminf_{x \to_K x_0} T_K(x) \quad .$$

■

We recall that when K is closed and convex, both the contingent cone $T_K(x_0)$ and the Clarke tangent cone $C_K(x_0)$ coincide with the tangent cone to K at x_0:

(20) $$C_K(x_0) = T_K(x_0) = \mathrm{cl}\left(\bigcup_{h>0} \frac{1}{h}(K-x_0)\right) \quad .$$

2. CONTINGENT AND CLARKE DERIVATIVE OF A SET-VALUED MAP

We adapt to the case of a set-valued map the intuitive definition of a derivative of a function in terms of the tangent to its graph.

Let F *be a strict set-valued map from* $K \subset X$ *to* Y. (We say that F is strict if its images F(x) are nonempty for all $x \in K$.) *Let* $x_0 \in K$ *and* $y_0 \in F(x_0)$.

We denote by $DF(x_0,y_0)$ *the set-valued map from* X *to* Y *whose graph is the contingent cone* $T_{\mathrm{graph}(F)}(x_0,y_0)$ *to the graph of* F *at* (x_0,y_0).

In other words,

(1) $$v_0 \in DF(x_0,y_0)(u_0) \text{ if and only if } (u_0,v_0) \in T_{\mathrm{graph}(F)}(x_0,y_0) \quad .$$

We observe that v_0 belongs to $DF(x_0,y_0)(u_0)$ if and only if

(2) $$\begin{cases} \text{There exist sequences } h_n \to 0+,\ u_n \to u_0 \text{ and } v_n \to v_0 \\ \text{such that } v_n \in \dfrac{F(x_0+h_nu_n) - y_0}{h_n} \text{ for all } n \quad . \end{cases}$$

Definition 1

We shall say that the set-valued map $DF(x_0,y_0)$ from X to Y is the "contingent derivative" of F at $x_0 \in K$ and $y_0 \in F(x_0)$.

It is a "process", i.e. a positively homogeneous set-valued map (since its graph is a cone) with closed graph.

We shall give an analytical characterization of $DF(x_0,y_0)$, which justifies that the above definition is a reasonable candidate for capturing the idea of a derivative as a (suitable) limit of differential quotients. We extend F to X by setting $F(x) = \phi$ when $x \notin K$.

Let F be a set-valued map from $K \subset X$ to Y and (x_0,y_0) belong to graph(F). Then v_0 belongs $DF(x_0,y_0)(u_0)$ if and only if

$$(3) \qquad \liminf_{\substack{h \to 0+ \\ u \to u_0}} d\left(v_0\,,\frac{F(x_0+hu) - y_0}{h}\right) = 0 \quad .$$

▲

When F is a single-valued map, we set

$$(4) \qquad DF(x_0,y_0) = DF(x_0) \quad ,$$

since $y_0 = F(x_0)$. The above formula shows that in this case, v_0 belongs to $DF(x_0)(u_0)$ if and only if

$$(5) \qquad \liminf_{\substack{h \to 0+ \\ u \to u_0}} \frac{\|F(x_0+hu) - F(x_0) - hv_0\|}{h} = 0 \quad .$$

If F is differentiable, then $DF(x_0)(u_0) = \nabla F(x_0)\cdot u_0$. When the graph of F is convex, we observe that v_0 belongs to $DF(x_0,y_0)(u_0)$ if and only if

$$(6) \qquad \liminf_{u \to u_0} \left(\inf_{h>0} d\left(v_0\,,\frac{F(x_0+hu) - y_0}{h}\right)\right) = 0 \quad .$$

We say that F is Lipschitz on a neighborhood $\mathcal{U}$ of $x_0 \in \mathrm{Int}\,(K)$ if there exists a constant $c > 0$ such that $\forall x,y \in \mathcal{U}$, $F(x) \subset F(y) + c\|x-y\|B$.

Proposition 1

Assume that F is Lipschitz on a neighborhood of x_0 (belonging to Int K). Then v_0 belongs to $DF(x_0,y_0)(u_0)$ if and only if

(7) $$\liminf_{h\to 0+} d\left(v_0 , \frac{F(x_0+hu_0) - y_0}{h}\right) = 0 \quad .$$

Furthermore, if the dimension of Y is finite, then

(8) $$\mathrm{Dom}\, DF(x_0,y_0) = X \quad .$$

Proof

a) The first statement follows from the fact that

(9) $$F(x_0+hu) - y_0 \subset F(x_0+hu_0) - y_0 + \ell h\|u-u_0\| B$$

when both h and $\|u-u_0\|$ are small.

b) Let u_0 belong to X. Then, for all $h > 0$ small enough,

(10) $$y_0 \in F(x_0) \subset F(x_0+hu_0) + \ell h\|u_0\| B \quad .$$

Hence, there exists $v_h \in F(x_0+hu_0)$ such that $(v_h-y_0)/h$ belongs to $\ell\|u_0\|B$, which is compact. A subsequence $(v_{h_n} - y_0)/h_n$ converges to some v_0, which belongs to $DF(x_0,y_0)(u_0)$. ∎

We point out that

(11) $$\forall x_0 \in K,\ \forall y_0 \in F(x_0),\quad DF(x_0,y_0)^{-1} = D(F^{-1})(y_0,x_0) \quad .$$

Indeed, to say that $(u_0,v_0) \in T_{\mathrm{graph}(F)}(x_0,y_0)$ amounts to saying that $(v_0,u_0) \in T_{\mathrm{graph}(F^{-1})}(y_0,x_0)$.

Contingent derivatives allow to "differentiate" restrictions of a map or a set-valued map to a subset.

Proposition 2

Let F be a single-valued map from an open subset Ω of X to Y differentiable at $x_0 \in K$ and K be a nonempty subset of Ω. Then

$$(12) \qquad DF|_K(x_0)u_0 = \begin{cases} \nabla F(x_0)u_0 & \text{if } u_0 \in T_K(x_0) \\ \phi & \text{if } u_0 \notin T_K(x_0) \end{cases}$$

▲

Proof

If F is a differentiable single-valued map at x_0 and u_0 belongs to $T_K(x_0)$, then there exist sequences $h_n \to 0$ and $u_n \to u_0$ such that $x_0 + h_n u_n$ belongs to K. Since

$$F|_K(x_0 + h_n u_n) = F(x_0 + h_n u_n) = F(x_0) + h_n(\nabla F(x_0)u_n + 0(h_n))$$

we deduce that the elements $v_n := \nabla F(x_0)u_n + 0(h_n)$ converge to $\nabla F(x_0)u_0$ and belong to $\left(F|_K(x_0 + h_n u_n) - F|_K(x_0)\right)/h_n$. Therefore,

$$DF|_K(x_0, F(x_0))(u_0) = \nabla F(x_0)u_0 \quad .$$

▲

We follow the same procedure for defining the Clarke derivative of a set-valued map from X to Y.

Let (x_0, y_0) belong to the graph of F. We denote by $CF(x_0, y_0)$ the closed convex process from X to Y whose graph is the Clarke tangent cone $C_{\text{graph}(F)}(x_0, y_0)$ to the graph of F at (x_0, y_0). Briefly

$$(13) \quad v_0 \in CF(x_0, y_0)(u_0) \text{ if and only if } (u_0, v_0) \in C_{\text{graph}(F)}(x_0, y_0).$$

Definition 2

We shall say that the closed convex process $CF(x_0, y_0)$ from X to Y is the Clarke derivative of F at $x_0 \in \text{Dom } F$ and $y_0 \in F(x_0)$.

We observe that v_0 belongs to $CF(x_0, y_0)(u_0)$ if and only if

$$(14)\qquad \begin{cases} \text{for all sequences of elements } (x_n, y_n, h_n) \in \text{graph}(F) \times \overset{\circ}{\mathbb{R}}_+ \\ \text{converging to } (x_0, y_0, 0), \text{ there exist sequences of} \\ \text{elements } u_n \text{ converging to } u_0 \text{ and } v_n \text{ converging to} \\ v_0 \text{ such that } y_n + h_n v_n \in F(x_n + h_n u_n) \text{ for all } n > 0 \end{cases}$$

We shall avoid using the analytical formula playing the role of formula (3), which involves a sort of mixture of both limsup and liminf. It is simplified when F is Lipschitz on a neighborhood of $x_0 \in \text{Int Dom } F$.

Proposition 3

Assume that F is Lipschitz on a neighborhood of an element $x_0 \in \text{Int Dom } F$. Then v_0 belongs to $CF(x_0, y_0)(u_0)$ if and only if

$$(15)\qquad \lim_{x \to_K x_0,\, h \to 0+} d\left(v_0, \frac{F(x + hu_0) - y_0}{h}\right) = 0 \quad .$$

▲

The proof is analogous to the one of Proposition 1. When F is single-valued, we shall set

$$(16)\qquad CF(x_0) := CF(x_0, F(x_0)) \quad .$$

If F is continuously differentiable at x_0, we have

$$(17)\qquad CF(x_0) = \nabla F(x_0) \quad .$$

Naturally, the formula on Clarke derivatives of inverses is obvious:

$$(18)\qquad \forall (x_0, y_0) \in \text{graph}(F),\ CF(x_0, y_0)^{-1} = C(F^{-1})(y_0, x_0) \quad .$$

Proposition 4

Let F be a single-valued map from an open subset Ω of X to Y, continuously differentiable at $x_0 \in \Omega$, and K be a nonempty subset of X. Then

$$(19) \qquad CF|_K(x_0)u_0 = \begin{cases} \nabla F(x_0)u_0 & \text{if} \quad u_0 \in C_K(x_0) \\ \phi & \text{if} \quad u_0 \notin C_K(x_0) \end{cases} \quad .$$

▲

Proof

Let $(x_n,h_n) \in K \times \overset{\circ}{\mathbb{R}}_+$ converge to $(x_0,0)$ in $K \times \overset{\circ}{\mathbb{R}}_+$. If u_0 belongs to $C_K(x_0)$, there exists a sequence of elements u_n converging to u_0 such that $x_n + h_n u_n$ belongs to K for all n. Then

$$F|_K(x_n + h_n u_n) = F(x_n + h_n u_n) = F(x_n) + h_n(\nabla F(x_n)u_n + 0(h_n)).$$

Since F is continuously differentiable, the sequence of elements $v_n := \nabla F(x_n)u_n + 0(h_n)$ converges to $\nabla F(x_0)u_0$ and we have $F|_K(x_n) + h_n v_n = F|_K(x + h_n u_n)$ for all n.

■

Proposition 5

Let F be a monotone map from X to X* and (x_0,y_0) belong to the graph of F. Then the Clarke derivative of F at (x_0,y_0) is a monotone (closed convex) process from X to X*.

▲

Proof

Let (u^i,v^i) (i=1,2) belong to the graph of the Clarke derivative $CF(x_0,y_0)$. We take $x_n := x_0$, $y_n := y_0$ and $h_n > 0$ converging to 0. Then there exist sequences (u_n^i,v_n^i) converging to (u^i,v^i) such that $y_0 + h_n v_n^i \in F(x_0 + h_n u_n^i)$ for all $n \geq 0$ $(i = 1,2)$. Since F is monotone, we deduce that

$$h_n^2 \langle v_n^1 - v_n^2 , u_n^1 - u_n^2 \rangle$$

$$= \langle y_0 + h_n v_n^1 - (y_0 + h_n v_n^2) , x_0 + h_n u_n^1 - (x_0 + h_n u_n^2) \rangle \geq 0 \quad .$$

■

An important class of monotone maps from X to X* is provided by the subdifferentials of convex functions.

Definition 3

Let $U : X \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function. Assume that U is subdifferentiable at x_0 and let $p_0 \in \partial U(x_0)$ be a subgradient of U at x_0. We shall say that

(20) $$\partial^2 U(x_0,p_0); = C\partial U(x_0,p_0)$$

is the second derivative of U at (x_0,p_0). ▲

Then $\partial^2 U(x_0,p_0)$ is a monotone closed convex process from X to X^*.

Proposition 6

Let V be a proper lower semicontinuous convex function from X to $\mathbb{R} \cup \{+\infty\}$ and V^* its conjugate function. Then

(21) $$\partial^2 V^*(p_0,x_0) = (\partial^2 V(x_0,p_0))^{-1} \quad .$$

Furthermore, if $q_0 \in \partial^2 V(x_0,p_0)(u_0)$, then

(22) i) $D_+V(x_0)(u_0) = \langle p_0,u_0 \rangle$ ii) $D_+V^*(p_0)(q_0) = \langle q_0,x_0 \rangle$.

where we set

$$D_+V(x_0)(u_0) := \liminf_{\substack{h \to 0+ \\ u \to u_0}} \left(\frac{V(x_0+hu) - V(x_0)}{h} \right) .$$

▲

Proof

a) The first equality is straightforward, since ∂V^* is the inverse of ∂V.

b) Since V is convex, upper contingent and Clarke derivatives coincide. Since $q_0 \in C(\partial V(x_0,p_0))(u_0)$, we can associate with a sequence of elements $h_n > 0$ converging to 0 a sequence of elements (u_n,q_n) converging to (u_0,q_0) such that $p_0 + h_n q_n$ belongs to $\partial V(x_0 + h_n u_n)$ for all n. We obtain the inequalities

i) $V(x_0+h_n u_n) - V(x_0) \leq h_n \langle p_0 + h_n q_n, u_n \rangle$

ii) $V(x_0) - V(x_0+h_n u_n) \leq -h_n \langle p_0,u_n \rangle$

from which we deduce formula (22)i). The second formula (22)ii) is obtained in the same way. ■

By taking for function V the indicator of a closed convex subset, we obtain:

Corollary 1

Let K be a nonempty closed convex subset, x_0 belong to K and p_0 belong to the normal cone $N_K(x_0)$ to K at x_0.

If q_0 belongs to $CN_K(x_0,p_0)(u_0)$, then

$$(23)\quad \begin{cases} \text{i)} & p_0 \in N_K(x_0),\ u_0 \in T_K(x_0) \text{ and } \langle p_0,u_0\rangle = 0 \\ \text{ii)} & D_+\sigma_K(p_0)(q_0) = \langle q_0,x_0\rangle \quad . \end{cases}$$

(where σ_K is the support function of K). ▲

Proposition 7

Let K be a nonempty closed convex subset of a Hilbert space X and p_0 belong to the normal cone $N_K(x_0)$. Let N_K denote the set-valued map $x \to N_K(x)$ and π_K denote the Lipschitz single-valued map associating to x its best approximation $\pi_K(x) \in K$ by elements of K. Then the two following statements are equivalent.

a) $q_0 \in CN_K(x_0,p_0)(u_0)$

b) $u_0 \in C\pi_K(x_0+p_0)(u_0+q_0)$.

The same result holds when the Clarke derivative is replaced by the contingent derivative. ▲

Proof

We recall that p belongs to the normal cone $N_K(x)$ if and only if $x = \pi_K(x+p)$.

a) Assume that q_0 belongs to $CN_K(x_0,p_0)(u_0)$. Let us consider a sequence of elements $(y_n,h_n) \in X \times \mathbb{R}_+^{\circ}$ converging to $(x_0+p_0,0)$. We set $x_n := \pi_K(y_n)$, which converges to $x_0 = \pi_K(x_0+p_0)$ and $p_n := y_n - x_n$, which converges to p_0. Then there exist sequences of elements u_n and q_n converging to u_0 and q_0 such that $p_n + h_n q_n$ belongs to $N_K(x_n + h_n u_n)$ for all n, i.e., such that $\pi_K(y_n) + h_n u_n = \pi_K(y_n + h_n(q_n + u_n))$ for all n. Hence u_0 belongs to $C\pi_K(x_0 + p_0)(u_0 + q_0)$.

b) Conversely, assume that u_0 belongs to $C\pi_K(x_0 + p_0)(u_0 + q_0)$. Let $(x_n, p_n, h_n) \in \text{graph } N_K \times \overset{\circ}{\mathbb{R}}_+$ converge to $(x_0, p_0, 0)$. Since $x_n + p_n$ converges to $x_0 + p_0$, there exist sequences of elements u_n and w_n converging to u_0 and $u_0 + q_0$ such that $x_n + h_n u_n = \pi_K(x_n + p_n) + h_n u_n = \pi_K(x_n + p_n + h_n w_n)$ for all n. Then $q_n := w_n - u_n$ converges to u_0 and we deduce that $p_n + h_n q_n \in N_K(x_n + h_n u_n)$ for all n. Hence q_0 belongs to $CN_K(x_0, p_0)(u_0)$.

■

Corollary 2

Let us consider the set-valued map associating to $x \in \mathbb{R}^n_+$ the normal cone $N_{\mathbb{R}^n_+}(x)$ to $\mathbb{R}^n_+$ at x. Let p^0 belong to $N_{\mathbb{R}^n_+}(x^0)$. Then q^0 belongs to $CN_{\mathbb{R}^n_+}(x^0, p^0)(u^0)$ if and only if

$$q_i^0 \in \begin{cases} \{0\} & \text{if } x_i^0 > 0 \quad (\text{and thus, } p_i^0 = 0) \\ \phi & \text{if } x_i^0 = 0 \quad p_i^0 \leq 0 \text{ and } u_i \neq 0 \\ \mathbb{R} & \text{if } x_i^0 = 0 \quad p_i^0 < 0 \text{ and } u_i = 0 \\ \{0\} & \text{if } x_i^0 = 0 \quad p_i^0 = 0 \text{ and } u_i = 0 \end{cases}$$

▲

Proof

We observe that $\pi_{\mathbb{R}^n_+}(x_1, \ldots, x_n) = (\pi(x_1), \ldots, \pi(x_n))$ where $\pi(x) = 0$ when $x \leq 0$ and $\pi(x) = x$ when $x \geq 0$. Since $C\pi(x)(u) = 0$ when $x < 0$, u when $x > 0$ and $C\pi(0)(u) = \phi$ when $u \neq 0$ and $C\pi(0)(0) = 0$, we obtain the above corollary.

■

3. THE INVERSE FUNCTION THEOREM

We shall extend the usual Inverse Function Theorem for continuously differentiable single-valued maps to the case of set-valued maps with closed graph.

We set

$$\underline{d}(A,B) := \sup_{x\in A} \inf_{y\in B} \|x-y\| \quad .$$

We make precise what we mean by Lipschitz behavior.

<u>Definition 1</u>

Let F be a proper set-valued map from X to Y and (x_0,y_0) belong to the graph of F. We say that F^{-1} is "pseudo-Lipschitz" around (y_0,x_0) if there exists a neighborhood $\mathcal{W}$ of y_0, two neighborhoods $\mathcal{U}$ and $\mathcal{V}$ of x_0, $\mathcal{U} \subset \mathcal{V}$, and a constant $\ell > 0$ such that

$$(1) \quad \begin{cases} \text{i)} & \forall y \in \mathcal{W}, \quad F^{-1}(y) \cap \mathcal{U} \neq \phi \\ \text{ii)} & \forall y_1,y_2 \in \mathcal{W}, \ \underline{d}(F^{-1}(y_1)\cap\mathcal{U},\ F^{-1}(y_2)\cap\mathcal{V}) \leq \ell\|y_1-y_2\| \ . \end{cases}$$

▲

See Rockafellar [8] for a study of these properties.

<u>Theorem 1</u>. Let F be a closed set-valued map from X to Y and (x_0,y_0) belong to graph (F). We assume that

$$(2) \quad \begin{cases} \text{i)} & \text{both X and Y are finite dimensional} \\ \text{ii)} & \text{the Clarke derivative } CF(x_0,y_0) \text{ of F at } (x_0,y_0) \\ & \text{is surjective (i.e., } \operatorname{Im} CF(x_0,y_0) = Y). \end{cases}$$

Then F^{-1} is pseudo-Lipschitz around (y_0,x_0). ▲

The same statement is proved in Rockafellar [8] by other methods, using the properties of Rockafellar [7].

Before proving this theorem, let us mention a useful corollary on nonlinear equations "with constraints".

(3) $\begin{cases}\text{Find a solution } x \text{ \underline{in K} to the equation } F(x) = y \text{ where}\\ \text{the solution is required to belong to a closed subset } K.\end{cases}$

Corollary 1

Let F be a map from an open subset Ω of X and K be a closed subset of X. Let $F|_K$ denote the restriction of F to K.

We posit the following assumptions

(4) $\begin{cases} \text{i)} & \text{the dimensions of X and Y are finite} \\ \text{ii)} & F \text{ is continuously differentiable at } x_0 \in \Omega \cap K \\ \text{iii)} & \forall v \in Y, \exists u \in C_K(x_0), \text{ the Clarke tangent cone} \\ & \text{to K at } x_0, \text{ solution to } \nabla F(x_0)u = v \,. \end{cases}$

Then

(5) F^{-1} is pseudo-Lipschitz around (y_0, x_0).

▲

When K is the whole space X, we obtain the usual Inverse Function Theorem for single-valued maps.

We shall deduce Theorem 1 from a rather general result, which is a slight modification of a theorem due to Lebourg [1] (see also Aubin [7]).

Theorem 2. Let F be a closed set-valued map from a Banach space X to a Banach space Y. Let $(x_0, y_0) \in \text{graph}(F)$ be fixed. We assume that there exist constants $\alpha \in [0,1[, \eta > 0$ and $c > 0$ such that

for all $(x,y) \in \text{graph}(F)$ satisfying $\|x-x_0\| + \|y-y_0\| \leq \eta$, for all $v \in Y$, there exist $u \in X$ and $w \in Y$ such that

$$(6)\quad \begin{cases} \text{i)} & v \in DF(x,y)(u) + w \\ \text{ii)} & \|u\| \le c\|v\| \quad \text{and} \quad \|w\| \le \alpha\|v\| \end{cases} .$$

Let us set

$$(7)\quad r: = \frac{\eta(1-\alpha)}{3(1+\alpha+c)} \quad , \quad F_0^{-1}(y): = F^{-1}(y) \cap (x_0 + \frac{c+2\alpha}{1-\alpha} rB)$$

$$F_1^{-1}(y) := F^{-1}(y) \cap \left(x_0 + \frac{3(c+2\alpha)}{1-\alpha} rB\right) \quad .$$

Then F^{-1} is pseudo-Lipschitz around (y_0, x_0): Namely,

$$(8)\quad \begin{cases} \text{i)} & \forall y \in y_0 + r\overset{\circ}{B} \quad , \quad F_0^{-1}(y) \neq \phi \\ \text{ii)} & \forall y_1, y_2 \in y_0 + r\overset{\circ}{B} \quad , \\ & \underline{d}(F_0^{-1}(y_1), F_1^{-1}(y_2)) \le \frac{c+2\alpha}{1-\alpha} \|y_1 - y_2\| \end{cases}$$

Proof

Let y_1 and y_2 belong to the open ball $y_0 + r\overset{\circ}{B}$. Assume for the time that there exists x_1 satisfying

$$(9)\quad x_1 \in F_0^{-1}(y_1): = F^{-1}(y_1) \cap (x_0 + \ell rB) \quad \text{where} \quad \ell: = \frac{c+2\alpha}{1-\alpha} \quad .$$

This is possible when we take $y_1 = y_0$ and $x_1 = x_0$! We associate with any $\rho \in]\|y_1-y_2\|, 2r]$ the number $\varepsilon: = \dfrac{\|y_1-y_2\|}{\|y_1-y_2\| + \ell\rho}$,

which satisfies

$$(10)\quad \frac{3\|y_1 - y_2\|}{2\eta} \le \varepsilon < \frac{1-\alpha}{1+c+\alpha} \quad .$$

We apply Ekeland's Theorem (see Ekeland [1] or Aubin [1], p.174) to the continuous function V defined on the graph of F by $V(x,y) = \|y_2 - y\|$.

Since it is complete, there exists $(\bar{x},\bar{y}) \in \text{graph}(F)$ such that

$$(11)\quad \begin{cases} \text{i)} & \|\bar{y}-y_2\| + \varepsilon(\|\bar{x}-x_1\| + \|\bar{y}-y_1\|) \leq \|y_1-y_2\| \\ \text{ii)} & \forall (x,y)\in \text{graph}(F),\ \|\bar{y}-y_2\| \leq \|y-y_2\| + \varepsilon(\|\bar{x}-x\|+\|\bar{y}-y\|) \ . \end{cases}$$

Inequality (11)i) implies that

$$\|\bar{x}-x_1\| + \|\bar{y}-y_1\| \leq \frac{1}{\varepsilon}\|y_1 - y_2\| \leq \frac{2\eta}{3} \quad .$$

Therefore

$$\|\bar{x}-x_0\| + \|\bar{y}-y_0\| \leq \frac{2\eta}{3} + \|x_0 - x_1\| + \|y_0 - y_1\|$$

$$\leq \frac{2\eta}{3} + \left(\frac{c+2\alpha}{1-\alpha} + 1\right) r = \frac{2\eta}{3} + \frac{1+\alpha+c}{1-\alpha} r = \frac{2\eta}{3} + \frac{\eta}{3} = \eta \quad .$$

Consequently, we can apply assumption (6) with $v: = y_2-\bar{y}$: there exist u and w satisfying

$$(12)\quad \begin{cases} \text{i)} & y_2-\bar{y} \in DF(\bar{x},\bar{y})(u) + w \\ \text{ii)} & \|u\| \leq c\|y_2-\bar{y}\| \text{ and } \|w\| \leq \alpha\|y_2-\bar{y}\| \quad . \end{cases}$$

By the very definition of the contingent derivative $DF(\bar{x},\bar{y})$, we can associate to any $\delta > 0$ elements $h \in]0,\delta]$, $u_\delta \in \delta B$ and $v_\delta \in \delta B$ such that the pair (x,y) defined by

$$x = \bar{x} + hu + hu_\delta \quad , \quad y = \bar{y} + h(y_2-\bar{y}) - hw - hv_\delta$$

belongs to the graph of F. Using this pair in inequality (11)ii), we obtain

$$\|y_2-\bar{y}\| \leq (1-h)\|y_2-\bar{y}\| + h\|w\| + h\varepsilon(\|u\| + \|y_2-\bar{y}\| + \|w\|)$$

$$+ h((1+\varepsilon)\|v_\delta\| + \varepsilon\|u_\delta\|) \quad .$$

We divide this inequality by $h > 0$ and we let δ converge to 0. We get

$$\|y_2-\bar{y}\| \le \varepsilon(\|u\| + \|y_2-\bar{y}\|) + (1+\varepsilon)\|w\| \quad .$$

Therefore, inequalities (12)ii) imply that

$$\|y_2-\bar{y}\| \le (\varepsilon(c+1) + \alpha(1+\varepsilon))\|y_2-\bar{y}\| \quad .$$

Since $\varepsilon < \frac{1-\alpha}{c+1+\alpha}$, we infer that $y_2 = \bar{y}$ and thus that $\bar{x}$ is a solution to the inclusion $y_2 \in F(\bar{x})$; by setting $y_2 = \bar{y}$ in inequality (11)i), we get

$$\|\bar{x}-x_1\| \le (\frac{1}{\varepsilon} - 1)\|y_1-y_2\| = \ell\rho \le 2\ell r \quad .$$

Therefore, $\bar{x}$ belongs to $F^{-1}(y_2) \cap (x_1 + 2\ell rB) \subset F_1^{-1}(y_2)$ and thus,

$$d(x_1,F_1^{-1}(y_2)) \le \|\bar{x}-x_1\| \le (\frac{1}{\varepsilon} - 1)\|y_1-y_2\| = \ell\rho \quad .$$

By letting ρ converge to $\|y_1-y_2\|$, we deduce that

$$(13) \qquad d(x_1,F_1^{-1}(y_2)) \le \ell\|y_1-y_2\| \quad .$$

We can always take $(x_1,y_1) = (x_0,y_0)$. We thus have proved that

$$(14) \qquad \forall y_2 \in y_0 + r\overset{o}{B},\ \exists x_2 \in F_0^{-1}(y_2) := F^{-1}(y_2) \cap \left(x_0 + \frac{c+2\alpha}{1-\alpha}\, rB\right)$$

(because $\|y_2-y_0\| < r$ instead of $2r$).

In other words, the set-valued map F_0^{-1} has nonempty images when y ranges over the open ball $y_0 + r\overset{o}{B}$. Inequality (13) implies that

$$\underline{d}(F_0^{-1}(y_1),F_1^{-1}(y_2)) := \sup_{x_1\in F_0^{-1}(y_1)} d(x_1,F_1^{-1}(y_2))$$

$$\le \frac{c+2\alpha}{1-\alpha}\|y_1 - y_2\| \quad .$$

■

The proof of the Inverse Function Theorem 1 follows from the above theorem and the following lemma.

Lemma 1

Let us assume that the spaces X and Y are finite dimensional. Let (x_0, y_0) belong to the graph of F. We assume that

(15) the Clarke derivative $CF(x_0, y_0)$ maps X onto Y.

Then, for all $\alpha > 0$ there exist constant $c > 0$ and $\eta > 0$ such that for all $(x,y) \in \text{graph}(F)$ satisfying

$$\|x-x_0\| + \|y-y_0\| \leq \eta$$

and for all $v \in Y$, there exist $u \in X$ and $w \in Y$ satisfying

(16) $$v \in DF(x,y)(u) \ , \quad \|u\| \leq c\|v\| \quad \text{and} \quad \|w\| \leq \alpha\|v\| \quad .$$

▲

Proof

Since $CF(x_0, y_0)$ is a closed convex process, Robinson-Ursescu's Theorem implies the existence of $\gamma > 0$ such that

(17) $$\gamma B \subset CF(x_0, y_0)(B) \quad .$$

(See Robinson [2], Ursescu [1].)

Let $$K = (B \times \gamma B) \cap \text{graph } CF(x_0, y_0) = (B \times \gamma B) \cap C_{\mathcal{F}}(x_0, y_0)$$

where $C_{\mathcal{F}}(x_0, y_0)$ is the Clarke tangent cone to the graph $\mathcal{F}$ of F at (x_0, y_0). The subset K is compact because the spaces X and Y are finite dimensional.

Since the Clarke tangent cone is the Liminf of the contingent cones:

$$C_{\mathcal{F}}(x_0, y_0) = \bigcap_{\alpha>0} \bigcup_{\eta>0} \bigcap_{(x,y) \in B_{\mathcal{F}}(x_0, y_0; \eta)} (T_{\mathcal{F}}(x,y) + \alpha(B \times B))$$

and since $K \subset C_F(x_0,y_0)$ is compact, we deduce that, for all $\alpha > 0$, there exists $\eta > 0$ such that for all $(u_0,v_0) \in K$, $(x,y) \in B_F(x_0,y_0;\eta)$, we have $(u_0,v_0) \in T_F(x,y) + \alpha(B \times B)$.

Now, take v in Y. Then $v_0 := \frac{\gamma v}{\|v\|}$ belongs to γB and by (17) there exists $u_0 \in B$ such that (u_0,v_0) belongs to K. Then, for all $(x,y) \in B_F(x_0,y_0;\eta)$, there exist $u_\alpha \in \alpha B$ and $v_\alpha \in \alpha B$ such that $(u_0 - u_\alpha, v_0 - v_\alpha) \in T_F(x,y)$, i.e. such that

$$v_0 \in DF(x,y)(u_0 - u_\alpha) + v_\alpha \quad .$$

We set $u := \frac{\|v\|}{\gamma}(u_0 - u_\alpha)$ and $w = \frac{\|v\|}{\gamma} v_\alpha$. Then $v \in DF(x,y)(u)$, $\|u\| \leq \frac{1+\alpha}{\gamma}\|v\|$ and $\|w\| \leq \frac{\alpha}{\gamma}\|v\|$.

■

4. CLARKE TANGENT CONE OF AN INVERSE IMAGE

The Inverse Function Theorem allows to prove the formula

(1) $$C_L(x_0) \cap \nabla A(x_0)^{-1} C_M(Ax_0) \subset C_{L \cap A^{-1}(M)}(x_0)$$

under conditions weaker than the ones in Rockafellar [4].

<u>Theorem 1</u>. Let X and Y be finite dimensional spaces, A be a continuously differentiable operator from an open subset $\Omega \subset X$ to Y, $L \subset X$ and $M \subset Y$ be closed subsets of X and Y respectively.

We assume that there exists $x_0 \in \Omega \cap L \cap A^{-1}(M)$ such that

(2) $$\nabla A(x_0) C_L(x_0) - C_M(Ax_0) = Y \quad .$$

Then

$$C_L(x_0) \cap \nabla A(x_0)^{-1} C_M(Ax_0) \subset C_{L \cap A^{-1}(M)}(x_0) \quad .$$

▲

The theorem follows from a series of lemmas. First, we introduce the set-valued map F defined by

(3) $$F(x) := \begin{cases} A(x) - M & \text{when } x \in L \\ \phi & \text{when } x \notin L \end{cases} \quad .$$

We observe that

$$F^{-1}(y) = \{x \in L \mid Ax \in M+y\}$$

$$F^{-1}(0) = L \cap A^{-1}(M) \quad .$$

Lemma 1

Let y_0 belong to $F(x_0)$. The following conditions are equivalent.

a) $v_0 \in CF(x_0,y_0)(u_0)$

b) $u_0 \in C_L(x_0)$ and $v_0 \in \nabla A(x_0)u_0 - C_M(Ax_0 - y_0)$.

Consequently, $CF(x_0,y_0)$ is surjective if and only if

$$\nabla A(x_0)C_L(x_0) - C_M(Ax_0 - y_0) = Y \quad .$$

▲

Proof

a) Let us prove that a) implies b). We take sequences $(x_n,z_n,h_n) \in L \times M \times \overset{\circ}{\mathbb{R}}_+$ converging to $(x_0,Ax_0-y_0,0)$. Then $y_n: = A(x_n) - z_n$ converges to y_0 and, by a) there exist sequences u_n and v_n converging to u_0 and v_0 such that $x_n + h_nu_n \in L$ and $A(x_n + h_nu_n) \in M + y_n + h_nv_n$ for all n. This implies that u_0 belongs to $C_L(x_0)$ and that $\nabla A(x_0)u_0 - v_0$ belongs to $C_M(Ax_0 - y_0)$ because $w_n: = A(x_n + h_nu_n) - A(x_n) - v_n$ converges to $\nabla A(x_0)u_0 - v_0$ and because $z_n + h_nw_n$ belongs to M for all n.

b) Conversely, let us show that a) follows from b). We take sequences $(x_n,y_n,h_n) \in \text{graph }(F) \times \overset{\circ}{\mathbb{R}}_+$ converging to $(x_0,y_0,0)$. There exists a sequence u_n converging to u_0 such that $x_n + h_nu_n$ belongs to L and, since $Ax_n - y_n$ converges to $Ax_0 - y_0$ in M, there exists a sequence of elements w_n converging to $\nabla A(x_0)u_0 - v_0$ and satisfying $Ax_n - y_n + h_nw_n \in M$ for all n.

Then the sequence of elements $v_n: = A(x_n + h_nu_n) - Ax_n - w_n$ converges to v_0 and satisfies $y_n + h_nv_n \in F(x_n + h_nu_n)$ for all n.

■

Lemma 2

We posit assumptions of Theorem 1. There exists a neighborhood $\mathcal{U}_0$ of x_0 in L such that

(4) $$\forall x \in \mathcal{U}_0 \quad , \quad d(x,F^{-1}(0)) \leq \ell d_M(Ax) \quad .$$

▲

Proof

We take $y_0 = 0$ and $x_0 \in F^{-1}(0)$. The Inverse Function Theorem implies that F^{-1} is pseudo-Lipschitz around $(0,x_0)$: then there exists a neighborhood $\mathcal{U}$ of x_0, a ball of radius r in Y and a constant $\ell > 0$ such that

$$\forall y \in r\mathring{B}, \quad \forall x \in F^{-1}(y) \cap \mathcal{U} \quad , \quad d(x,F^{-1}(0)) \leq \ell\|y\| \quad .$$

We can choose $\mathcal{U}$ so small that $\|A(x) - A(x_0)\| < r$ when x ranges over $\mathcal{U}$. Any $x \in L \cap \mathcal{U}$ belongs to $F^{-1}(A(x) - \pi_M(Ax))$ because

$$\|A(x) - \pi_M(Ax)\| \leq \|Ax - Ax_0\| < r \quad .$$

Therefore we know that for all $x \in \mathcal{U}_0 := L \cap \mathcal{U}$,

$$d(x,F^{-1}(0)) \leq \ell\|A(x) - \pi_M(Ax) - 0\| = \ell d_M(Ax) \quad .$$

■

Proof of Theorem 1

Let u_0 belong to $C_L(x_0) \cap \nabla A(x_0)^{-1} C_M(Ax_0)$. There exist $\alpha > 0$ and $\beta > 0$ such that $x + hu_0$ belongs to $\mathcal{U}_0 := L \cap (x_0 + rB)$ when $\|x-x_0\| \leq \alpha$ and $h \leq \beta$. Since $F^{-1}(0) = L \cap A^{-1}(M)$, we deduce from the above lemma that

$$\frac{d(x+hu_0,F^{-1}(0))}{h} \leq c \frac{d_M(A(x+hu_0))}{h}$$

$$\leq c \frac{d_M(Ax+h\nabla A(x_0)u_0)}{h} + c \frac{\|A(x+hu_0) - A(x) - h\nabla A(x_0)u_0\|}{h} \quad .$$

The first term on the right-hand side converges to 0 because $\nabla A(x_0)u_0$ belongs to $C_M(Ax_0)$ and the second converges also to 0 because A is continuously differentiable. Hence u_0 belongs to $C_{F^{-1}(0)}(u_0)$.

■

5. REGULARITY OF SOLUTIONS TO CONVEX OPTIMIZATION PROBLEMS

We introduce

(1)
- i) two finite dimensional spaces X and Y
- ii) a linear operator A_0 from X to Y
- iii) two proper lower semicontinuous convex functions $U: X \to \mathbb{R}\cup\{+\infty\}$ and $V: Y \to \mathbb{R}\cup\{+\infty\}$.

We take

(2)
- i) $y_0 \in \text{Int}\,(\text{Dom}\,V - A_0\,\text{Dom}\,U)$
- ii) $p_0 \in \text{Int}\,(A_0^*\,\text{Dom}\,V^* + \text{Dom}\,U^*)$

We recall that the solutions $(x_0, q_0) \in X \times Y^*$ of the optimization problem

(3)
- i) $$U(x_0) + V(A_0x_0 + y_0) - \langle p_0, x_0\rangle = \min_{x\in X}\,(U(x) + V(A_0x + y_0) - \langle p_0, x\rangle)$$
- ii) $$U^*(-A_0^*q_0 + p_0) + V^*(q_0) - \langle q_0, y_0\rangle = \min_{q\in Y^*}\,(U^*(-A_0^*q + p_0) + V^*(q) - \langle q, y_0\rangle)$$
- iii) $$U(x_0) + U^*(-A_0^*q_0+p_0) + V(A_0x_0+y_0) + V^*(q_0) = \langle p_0, x_0\rangle + \langle q_0, y_0\rangle$$

are the solutions to the system of inclusions

(4)
- i) $p_0 \in \partial U(x_0) + A_0^*q_0$
- ii) $y_0 \in -A_0x_0 + \partial V^*(q_0)$.

(See Aubin [2], chapter 10 or Aubin [3], chapter 14.)

We shall study the behavior of the solutions (x_0, q_0) to this system with respect to the parameters p_0, y_0 and A_0.

For that purpose, let us denote by $F^{-1}(p,y,A)$ the subset of solutions (x,q) to the problem

$$(5) \quad \begin{cases} \text{i)} & p \in \partial U(x) + A^* q \\ \text{ii)} & y \in -Ax + \partial V^*(q) \end{cases} .$$

Let us recall that we denote by $\partial^2 U(x,p)$ the Clarke derivative of the set-valued map $x \to \partial U(x)$ at (x,p) where $p \in \partial U(x)$.

<u>Theorem 1</u>. We posit assumptions (1) and (2). Let (x_0, q_0) be a solution of problem (3). We assume that the monotone closed convex process from $X \times Y$ to itself defined by

$$\begin{pmatrix} \partial^2 U(x_0, p_0 - A_0^* q_0) & A_0^* \\ -A_0 & \left(\partial^2 V(Ax_0 + y_0, q_0)\right)^{-1} \end{pmatrix}$$

<u>is surjective</u>.

Then

$$(6) \qquad F^{-1} \text{ is pseudo-Lipschitz around } (p_0, y_0, A_0, x_0, q_0).$$

Furthermore, the derivative of F^{-1} is defined by the formula

$$(7) \qquad (\delta x, \delta q) \in CF^{-1}(p_0, y_0, A_0;\ x_0, q_0)(\delta p, \delta y, \delta A)$$

if and only if

$$(8) \quad \begin{pmatrix} \delta x \\ \delta q \end{pmatrix} \in \begin{pmatrix} \partial^2 U(x_0, p_0 - A_0^* q_0) & A_0^* \\ -A_0 & (\partial^2 V(Ax_0 + y_0, q_0))^{-1} \end{pmatrix}^{-1} \begin{pmatrix} \delta p - \delta A^* \cdot q_0 \\ \delta y + \delta A \cdot x_0 \end{pmatrix}$$

For simplicity, we set $G(x) := \partial U(x)$, $H(y) = \partial V(y)$ so that $H^{-1}(q) = \partial V^*(q)$.

Let F be the map from $X \times Y^*$ to $X^* \times Y \times L(X,Y)$ defined by

$$(9) \qquad (p,y,A) \in F(x,q)$$

if and only if

$$(10) \qquad \begin{cases} \text{i)} & p \in G(x) + A^* q \\ \text{ii)} & y \in -Ax + H^{-1}(q) \end{cases} .$$

We shall characterize the Clarke derivative of F in terms of the Clarke derivatives of the set-valued maps G and H (or H^{-1}) respectively.

<u>Lemma 1</u>

Let x_0, q_0 be a solution to the system of inclusions

$$(11) \qquad \begin{cases} \text{i)} & p_0 \in G(x_0) + A_0^* q_0 \\ \text{ii)} & y_0 \in -A_0 x_0 + H^{-1}(q_0) \end{cases} .$$

The following conditions are equivalent

$$(12)\ \text{a)} \quad (\delta p, \delta y, \delta A) \in CF(x_0, q_0; p_0, y_0, A_0)(\delta x, \delta q)$$

$$(12)\ \text{b)} \quad \begin{cases} \text{i)} & \delta p - \delta A^* \cdot q_0 \in CG(x_0, p_0 - A_0^* q_0)(\delta x) + A_0^* \delta q \\ \text{ii)} & \delta y + \delta A \cdot x_0 \in -A_0 \delta x + CH^{-1}(q_0, y_0 + A_0 x_0)(\delta q) \end{cases} .$$

▲

<u>Proof</u>

a) We prove that (a) implies (b). We choose sequences $(x_n, q_n, p_n', y_n', h_n)$ converging to $(x_0, q_0, p_0 - A_0^* q_0, y_0 + A_0 x_0, 0)$. By setting $A_n := A_0$, $p_n := p_n' + A_0^* q_n$ and $y_n := y_n' - A_0 x_n$, we see that $(x_n, q_n, p_n, y_n, A_n, h_n)$ converges to $(x_0, q_0, p_0, y_0, A_0, 0)$.

Therefore, by (a), there exist sequences of elements δx_n, δq_n, δp_n, δy_n and δA_n converging to δx, δq, δp, δy and δA such that

$$\begin{cases} \text{i)} \quad p_n' + h_n(\delta p_n - A_0^*\delta q_n - \delta A_n^* q_n - h_n \delta A_n^* \delta q_n) \in G\,(x_n + h_n \delta x_n) \\ \text{ii)} \quad y_n' + h_n(\delta y_n + A_0 \delta x_n + \delta A_n x_n + h_n \delta A_n \delta x_n) \in H^{-1}(q_n + h_n \delta q_n) \;. \end{cases}$$

Hence the system of inclusions (b) holds true.

b) Conversely, let us consider the system of inclusions (b) and let us prove (a). We choose sequences $(x_n, q_n, y_n, p_n, A_n, h_n)$ converging to $(x_0, q_0, y_0, p_0, A_0, 0)$. Then we know that there exist sequences of elements δx_n, δq_n, u_n and v_n converging to δx, δq, $\delta p - \delta A^* q_0 - A_0^* \delta q$, $\delta y + \delta A x_0 + A_0 \delta x$ respectively. We set

i) $\delta A_n := \delta A$, which converges to δA

ii) $\delta p_n := u_n + \delta A^* \cdot q_n + A_n^* \delta q_n + h_n \delta A^* \delta q_n$, which converges to δp

iii) $\delta y_n := v_n - \delta A x_n - A_n \delta x_n - h_n \delta A \delta x_n$, which converges to δx.

Hence

$$(p_n + h_n \delta p_n, y_n + h_n \delta y_n, A_n + h_n \delta A) \in F(x_n + h_n \delta x_n, q_n + h \delta q_n)$$

and consequently, inclusion a) holds true. ■

Hence Theorem 1 follows from the Inverse Function Theorem and the following Corollary.

<u>Corollary 1</u>

The closed convex process $CF(x_0, q_0; p_0, y_0, A_0)$ is surjective if and only if the closed convex process

$$\begin{pmatrix} CG(x_0, p_0 - A_0^* q_0) & A_0^* \\ -A_0 & (CH(A_0 x_0 + y_0, q_0))^{-1} \end{pmatrix}$$

is surjective.

▲

Remark

By eliminating q in the system (10), the solutions x are the solutions to the inclusion

$$(13) \qquad p \in G(x) + A^* H(Ax+y) \quad .$$

For studying the behavior of the solutions x to (13) with respect to the parameters p, y and A, we introduce the set-valued map E from X to $X^* \times Y \times L(X,Y)$ defined by

$$E(x) = \{(p,y,A) \,|\, p \in G(x) + A^* H(Ax+y)\} \quad .$$

We are tempted to use the Inverse Function Theorem. Unfortunately, we cannot express the Clarke derivative of E in terms of the Clarke derivatives of G and H and check its surjectivity under reasonable conditions.

Then, even if we are interested only in problem (13), we have to introduce an auxiliary variable q and replace the inclusion (13) by the equivalent system (10).

■

Example 1

We consider a minimization problem with equality constraints, defined by

(14)
- i) two finite dimensional spaces X and Y
- ii) a linear operator A_0 from X to Y
- iii) a lower semicontinuous convex function U from X to $\mathbb{R}$.

We take

(15)
$$\begin{cases} \text{i)} & y_0 \in -\mathrm{Int}\,(A_0\,\mathrm{Dom}\,U) \\ \text{ii)} & 0 \in \mathrm{Int}\,(\mathrm{Im}\,A_0^* + \mathrm{Dom}\,U^*) \end{cases} .$$

Let x_0 be a solution to the minimization problem

(16)
$$\begin{cases} \text{i)} & A_0 x_0 = -y_0 \\ \text{ii)} & U(x_0) = \min\limits_{A_0 x = -y_0} U(x) \end{cases}$$

and q_0 an associated Lagrange multiplier. Assume that

(17)
$$\begin{cases} \text{i)} & A_0 \text{ is surjective} \\ \text{ii)} & U \text{ is twice continuously differentiable} \\ & \text{at } x_0 \text{ and } \nabla^2 U(x_0) \text{ is positive-definite.} \end{cases}$$

Then F^{-1} is pseudo-Lipschitz around $(0, y_0, A_0, x_0, q_0)$. We set

(18)
$$\begin{cases} \text{i)} & J(x_0) = (A_0 \nabla^2 U(x_0)^{-1} A_0^*)^{-1} \\ \text{ii)} & A_0^+ = \nabla^2 U(x_0)^{-1} A_0^* J(x_0), \text{ which is a right} \\ & \text{inverse of } A_0 \\ \text{iii)} & q\otimes^* : A \in L(X,Y) \to A^* q \in X^* \\ \text{iv)} & x\otimes : A \in L(X,Y) \to Ax \in Y \end{cases} .$$

Then the Clarke derivative of the map F^{-1} is given by the formula:

$$\begin{pmatrix} \delta x \\ \delta q \end{pmatrix} = \begin{pmatrix} (1 - A_0^+ A_0)\nabla^2 U(x_0)^{-1} & -A_0^+ & -q_0\otimes^* \\ (A_0^+)^* & J(x_0) & x_0\otimes \end{pmatrix} \begin{pmatrix} \delta p \\ \delta y \\ \delta A \end{pmatrix}$$

Proof

We apply Theorem 1 to the case when V is the indicator of $\{0\}$. Then $\partial^2 V^*(q_0, A_0x_0 + y_0) = \partial^2 V^*(q_0, 0)$ is the constant map equal to 0. So inclusion (8) can be written

$$\begin{pmatrix}\delta x\\ \delta q\end{pmatrix} = \begin{pmatrix}\nabla^2 U(x_0) & A_0^*\\ -A_0 & 0\end{pmatrix}^{-1}\begin{pmatrix}\delta p - \delta A^* q_0\\ \delta y + \delta A x_0\end{pmatrix}$$

$$= \begin{pmatrix}(1-A_0^+A_0)\nabla^2 U(x_0)^{-1} & -A_0^+\\ (A_0^+)^* & J(x_0)\end{pmatrix}\begin{pmatrix}\delta p - \delta A^* q_0\\ \delta y + \delta A x_0\end{pmatrix}$$

which yields above formula.

Example 2

We consider the items defined by (14). We set $Y := \mathbb{R}^n$ and we take

$$(19)\quad \begin{cases} \text{i)} & y_0 \in -A_0 \operatorname{Dom} U - \overset{0}{\mathbb{R}}{}^n_+ \\ \text{ii)} & 0 \in \operatorname{Int}\,(A_0^* \mathbb{R}^n_+ + \operatorname{Dom} U^*) \end{cases} .$$

Let x_0 be a solution to the minimization problem

$$(20)\quad \begin{cases} \text{i)} & A_0 x_0 + y_0 \leq 0 \\ \text{ii)} & U(x_0) = \min_{A_0 x + y_0 \leq 0} U(x) \end{cases}$$

and $q_0 \geq 0$ be an associated Lagrange multiplier.

We denote by I_1 the set of indexes such that $(A_0x_0 + y_0)_i = 0$. We posit assumption (17) and

$$(21)\qquad \forall i \in I_1 \quad q_{0i} > 0 \quad .$$

Then F^{-1} is pseudo-Lipschitz around $(0, y_0, A_0, x_0, q_0)$. We write

$$(22)\quad \begin{cases} \text{i)} & \mathbb{R}^n = \mathbb{R}^{I_1} \times \mathbb{R}^{I_2} \quad \text{where } I_2 = \{i=1,\dots,n \mid i \notin I_1\} \\ \text{ii)} & q = (q^1,q^2) \quad , \quad A = (A^1,A^2) \\ \text{iii)} & J_1(x_0) = (A_0^1 \nabla^2 U(x_0)^{-1} A_0^{1*})^{-1} \\ \text{iv)} & A_0^{1+} = \nabla^2 U(x_0)^{-1} A_0^{1*} J_1(x_0) \end{cases}$$

Then the Clarke derivative of the set-valued map F^{-1} at $(0,y_0,A_0)$ is given by the inclusion written symbolically

$$\begin{pmatrix} \delta x \\ \delta q_1 \\ \delta q_2 \end{pmatrix} \in \begin{pmatrix} (1 - A_0^{1+} A_0^1)\nabla^2 U(x_0)^{-1} & -A_0^{1+} & -q_0^1 \otimes^* \\ (A_0^{1+})^* & J_1(x_0) & x_0^1 \otimes \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \delta p \\ \delta y_1 \\ \delta A_0^1 \end{pmatrix}$$

▲

<u>Proof</u>

We apply Theorem 1 to the case when V is the indicator $\psi_{-\mathbb{R}^n_+}$ of the cone $-\mathbb{R}^n_+$. Then $V^* = \psi_{\mathbb{R}^n_+}$ is the indicator of $\mathbb{R}^n_+$ and $\partial V^* = N_{\mathbb{R}^n_+}$ is the normal cone to $\mathbb{R}^n_+$. We take

$$-y_0 \in \text{Int } (A_0 \,\text{Dom}\, U + \mathbb{R}^n_+) = A_0 \,\text{Dom}\, U + \overset{\circ}{\mathbb{R}}{}^n_+$$

(which is the Slater condition).

Then (x_0,q_0) is a solution to the inclusion

$$(23)\quad \begin{cases} \text{i)} & 0 = \nabla U(x_0) + A_0^* q_0 \\ \text{ii)} & y_0 \in -A_0 x_0 + N_{\mathbb{R}^n_+}(q_0) \end{cases} \quad .$$

The latter condition implies that

$$(24)\quad \langle q_0, A_0 x_0 + y_0 \rangle = 0 \quad .$$

Since $A_0x_0 + y_0 \in -\mathbb{R}^n_+$, we deduce that

(25) if $(A_0x_0 + y_0)_i < 0$ then $q_{0i} = 0$.

Then q_0^2 is equal to 0 and we assumed that $q_{0i} > 0$ for all $i \in I_1$.

By Corollary 2.2, an element δy of $CN_{\mathbb{R}^n_+}(q_0, A_0x_0 + y_0)(\delta q)$ is defined by

for $i \in I_1$, $\delta y_i = 0$ and δq_i is arbitrary

for $i \in I_2$, $\delta y_i \in \mathbb{R}$ and δq_i is equal to 0.

Let us write $\mathbb{R}^n = \mathbb{R}^{I_1} \times \mathbb{R}^{I_2}$ and $q = (q^1, q^2)$. The domain of $\partial^2 V^*(q_0, A_0x_0 + y_0)$ is $\mathbb{R}^{I_1} \times \{0\}$ and $\partial V^*(q_0, A_0x_0 + y_0)(\partial q_1, 0) = \{0\} \times \mathbb{R}^{I_2}$. Hence the matrix of second derivatives can be written symbolically

$$\begin{pmatrix} \delta p \\ \delta y_1 \\ \delta y_2 \end{pmatrix} \in \begin{pmatrix} \nabla^2 U(x_0) & A_0^{1*} & A_0^{2*} \\ -A_0^1 & 0 & 0 \\ -A_0^2 & 0 & \mathbb{R} \end{pmatrix} \begin{pmatrix} \delta x \\ \delta q_1 \\ 0 \end{pmatrix}$$

Then it is surjective if and only of the matrix of linear operators

$$\begin{pmatrix} \nabla^2 U(x_0), & A_0^{1*} \\ -A_0^1 & 0 \end{pmatrix}$$

from $X \times \mathbb{R}^{I_1}$ to itself is surjective. This is the case by assumption (17). We can even invert the above inclusion explicitly and we obtain the formula for the Clarke derivative of the map F^{-1}. ∎

Example 3

We still consider the items defined by (14) and we consider a closed convex subset P of Y.

We take

$$(26)\qquad \begin{cases} y_0 \in \operatorname{Int}\ (P - A_0 \operatorname{Dom} U) \\ 0 \in \operatorname{Int}\ (A_0^* b(P) + \operatorname{Dom} U^*) \end{cases}$$

where $b(P) := \{q \mid \sup_{y \in P} \langle q, y\rangle < +\infty\}$ is the barrier cone of P. Let x_0 be a solution to the minimization problem

$$(27)\qquad \begin{cases} \text{i)} & Ax_0 \in P - y_0 \\ \text{ii)} & U(x_0) = \min\limits_{Ax \in P - y_0} U(x) \end{cases}$$

and q_0 an associated Lagrange multiplier. We posit assumption (17) and the following assumption on P.

$$(28)\qquad \begin{cases} \forall y \in Y,\ \exists q \text{ solution to} \\ y - J(x_0)q \in C\pi_P(Ax_0 + y_0 + q_0)(y + (1 - J(x_0))q) \end{cases} .$$

Then the conclusion of Theorem 1 holds true.

Proof

It is sufficient to check the surjectivity of

$$\begin{pmatrix} \nabla^2 U(x_0) & A_0^* \\ -A_0 & CN_P(Ax_0 + y_0, p_0)^{-1} \end{pmatrix}$$

This follows from Proposition 2.7. ■

This method, the use of the Inverse Function Theorem for set-valued maps, can be used to treat more general convex minimization problems.

6. SURJECTIVITY OF CLOSED CONVEX PROCESSES

We denote by K^-, the negative polar cone of a subset $K \subset X$, the closed convex cone of X^* defined by

$$(1) \qquad K^- = \{p \in X^* \mid \forall x \in K , \langle p,x\rangle \leq 0\} \quad .$$

We set $K^+ = -K^-$ and we recall that K is a closed convex cone if and only if $K = K^{--}$.

Let F be a closed convex process from X to Y (see Rockafellar [2]).

We define the set-valued map F^* from Y^* to X^* by

$$(2) \quad p \in F^*(q) \Longleftrightarrow \forall (x,y) \in \text{graph } (F), \ \langle p,x\rangle - \langle q,y\rangle \leq 0 \quad .$$

This amounts to saying that

$$(3) \quad p \in F^*(q) \Longleftrightarrow (p,-q) \text{ belongs to } \left(\text{graph } (F)\right)^- .$$

Therefore, F^* is also a closed convex process.

<u>Definition 1</u>

We shall say that the closed convex process F^* is the transpose of F. ▲

We observe that $F = F^{**}$. Since we are interested by the surjectivity of F, we shall extend to closed convex processes the theorem stating that a continuous linear operator is surjective if and only if its image is closed and its transpose is injective.

<u>Proposition 1</u>

A closed convex process F from X to Y is surjective if and only if its image is closed and $F^{*-1}\{0\} = 0$. ▲

<u>Proof</u>

a) We begin by proving that

(4) $$(\mathrm{Im}\,F)^{+} = F^{*-1}(0) \quad .$$

Indeed, q belongs to $(\mathrm{Im}\,F)^{+}$ if and only if

$$\forall (x,y) \in \mathrm{graph}(F) \quad , \quad <q,y> \geq 0$$

i.e., if and only if

$$\forall (x,y) \in \mathrm{graph}(F) \quad , \quad <0,x> - <q,y> \leq 0 \quad .$$

This amounts to saying that (0,q) belongs to the graph of F^{*}, i.e., that $q \in F^{*-1}(0)$.

b) The image of F is closed if and only if $\mathrm{Im}\,F = (F^{*-1}(0))^{+}$. The latter set is equal to Y if and only if $F^{*-1}(0) = \{0\}$. ■

We shall use this result to prove an extension of the Lax-Milgram Theorem to closed convex processes.

Definition 2

We shall say that a set-valued map F from X to X^{*} is X-elliptic if

(5) $$\begin{cases} \exists\, c > 0 \text{ such that, for any two } (x^{i},y^{i}) \in \mathrm{graph}(F),\ i=1,2, \\ \quad <y^{1} - y^{2}, x^{1} - x^{2}> \geq c\|x^{1} - x^{2}\|^{2} \quad . \end{cases}$$

▲

Lemma 1

The image of an X-elliptic map F with closed graph is closed and its inverse is single-valued and Lipschitz with constant c^{-1}.

Proof

The fact that F^{-1} is single-valued follows from (5) by taking $y = y^{1} = y^{2}$, and x_1, x_2 in $F^{-1}(y)$.

Inequality (5) implies also that

$$c\|F^{-1}(y^1) - F^{-1}(y^2)\|^2 \leq \|y^1 - y^2\|\|F^{-1}(y^1) - F^{-1}(y^2)\| \quad .$$

For proving that Im(F) is closed, let us consider a Cauchy sequence of elements $p_n \in \operatorname{Im} F$. Let us take x_n in $F^{-1}(p_n)$. Since F is X-elliptic, we deduce that

$$c\|x_n - x_m\|^2 \leq \langle p_n - p_m , x_n - x_m \rangle \leq \|x_n - x_m\|\|p_n - p_m\|$$

and therefore, that the sequence of elements x_n is a Cauchy sequence. Then the sequence of elements $(x_n,p_n) \in \operatorname{graph}(F)$ converges to some (x,p), which belongs to the graph of F since the latter is closed. Hence p belongs to Im(F). We have proved that it is complete, and thus, closed. ■

We deduce a surjectivity criterion analogous to the Lax-Milgram Theorem on X-elliptic continuous linear operators.

Proposition 2

Let F be an X-elliptic closed convex process from X to X^*. If $(\operatorname{Dom} F)^- \subset \operatorname{Im} F$ and if the domain of F is closed, then F is surjective and its inverse is a single-valued Lipschitz map from X^* to X.

▲

Proof

By assumption, $F^{*-1}(0) = \operatorname{Im} F^+$ (by (4)) is contained in $-(\operatorname{Dom} F)^{--} = -\operatorname{Dom} F$ since the domain of F is closed. Let us pick $x_0 \in F^{*-1}(0)$ and choose $y_0 \in F(-x_0)$. Since $(0,-x_0)$ belongs to graph $(F)^-$, we deduce that $\langle 0,x_0\rangle - \langle x_0,y_0\rangle \leq 0$. Since F is an X-elliptic process, we deduce that

$$c\|x_0\|^2 = c\|-x_0-0\|^2 \leq \langle -x_0-0,y_0-0\rangle = -\langle x_0,y_0\rangle \leq 0 \quad .$$

Hence $x_0 = 0$. Therefore, $F^{*-1}(0)$ is equal to $\{0\}$. Since Im F is closed by Lemma 1, Proposition 1 implies that F is surjective.

■

Corollary 1

Any X-elliptic closed convex process F whose domain is X is surjective and F^{-1} is a single-valued Lipschitz map from X to X^*. ▲

Corollary 2 (Lax-Milgram)

Any X-elliptic continuous linear operator from X to X^* is an isomorphism. ▲

REFERENCES

Aubin, Jean-Pierre

[1] Applied Abstract Analysis. Wiley. Interscience (1977).

[2] Applied Functional Analysis. Wiley. Interscience (1979).

[3] Mathematical Methods of Game and Economic Theory. North-Holland (1979).

[4] Gradients généralisés de Clarke. Ann. Sc. Math. Québec 2(1978):197-252.

[5] Further properties of Lagrange multipliers in nonsmooth optimization. Appl. Math. Optim. 6(1980):79-90.

[6] Derivatives and codifferentials of maps with closed convex graphs and convex operators. M.R.C. Tech. Rep. # 2011 (1979).

[7] Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions. In Mathematical Analysis and Applications, Advances in Mathematics, Suppl. Studies. Vol. 7A(1981): 159-229.

Aubin, J.-P. and F.H. Clarke

[1] Multiplicateurs de Lagrange en optimisation non convexe et applications. C.R.A.S. 285(1977):451-454.

[2] Monotone invariant solutions to differential inclusions. J. London Math. Soc. 16(1977):357-366.

Aubin, J.P. and I. Ekeland

[1] Applied Nonlinear Analysis. Wiley. Interscience. To appear.

Clarke, F.H.

[1] Generalized gradients and applications. Trans. Am. Math. Soc. 205(1975):247-262.

[2] A new approach to Lagrange multipliers. Math. Op. Rex. 1(1976):165-174.

[3] On the Inverse Function Theorem. Pacific J. Math. 64(1976):97-102.

[4] Book to appear.

Cornet, B.

[1] Cône tangent et cône normal à un ensemble. In thesis. Université de Paris-Dauphine (1981).

[2] Existence of Slow solutions for a class of differential inclusions. To appear.

Cornet, B. and G. Laroque

[1] Lipschitz properties of constrained demand functions and constrained maximizers. To appear.

Crouzeix, J-P.

[1] A relationship between the second derivatives of a convex function and of its conjugate. Math. Programming. 13 (1977): 364-365.

Ekeland, I.

[1] On the Variational Principle. J. Math. Anal. Appl. 47(1974):324-353.

Hiriart-Urruty, J.-B.

[1] Tangent cones, generalized gradients and mathematical programming in Banach spaces. Math. Op. Res. (1978)

Ioffe, A.D.

[1] Différentielles généralisées d'applications localement Lipschitziennes d'un espace de Banach dans un autre. C.R.A.S. 289(1979):637-640.

[2] Non smooth analysis and the theory of Fans. In Convex Analysis and Applications. R.B. Vinter Ed. To appear.

[3] Non smooth analysis: differential calculus of non differential mappings. To appear.

[4] Sous-différentielles approchées de fonctions numériques. C.R.A.S. (1981).

Lament, P-J.

[1] Approximation et optimisation. Hermann (1972).

Lebourg, G.

[1] Thesis. To appear.

Penot,

[1] A characterization of tangential regularity.

Robinson, S.M.

[1] Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear algorithms. Math. Prog. 7(1974): 1-16.

[2] Regularity and stability for convex multivalued functions. Math. Op. Res. 1(1976):130-143.

[3] Stability theory for systems of inequalities, Part I: Linear systems. SIAM Journal on Num. Anal. 12(1975): 754-769.

[4] First order conditions for general nonlinear optimization. SIAM Journal on Appl. Math. 30(1976):597-607.

Robinson, S.M.

[5] Stability theory for systems of inequalities, Part II: Differentiable nonlinear systems. SIAM Journal on Num. Anal. 13(1976):497-513.

[6] Generalized equations and their solutions, Part I: Basic theory. Math. Prog. Study 10(1979):128-141.

[7] Strongly regular generalized equations. Math. of O.R. 5(1980):43-62.

[8] Some continuity properties of polyhedral multifunctions. Tech. Summary Report No. 2014, Mathematics Research Center, University of Wisconsin-Madison (1979); Math. Progr. Studies. 14(1981):206-214.

[9] Generalized equations and their solutions, Part II: Applications to nonlinear programming (to appear).

Rockafellar, R.T.

[1] Convex Analysis. Princeton University Press. (1970)

[2] Monotone processes of convex and concave type. Am. Math. Soc. Memoir no. 77 (1967).

[3] La théorie des sous-gradients et ses applications à l'optimisation. Presses de l'Université de Montréal. (1978).

[4] Clarke's tangent cones and the boundaries of closed sets in $\mathbb{R}^n$. Nonlinear Analysis 3(1979):145-154.

[5] Directionally Lipschitzian functions and subdifferential calculus. Proc. London Math. Soc. 3(1979):331-356.

[6] Generalized directional derivatives and subgradients of non convex functions. Can. J. Math. 2(1980):257-280.

[7] Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming. Math. Prog. Study no. 5. R. Wets. Ed. To appear.

[8] Lipschitzian properties of multifunctions. To appear.

Ursescu, G.

[1] Multifunctions with closed convex graph. Czechs. Math. Journal 25(1975):438-49.