NONSMOOTH ANALYSIS

Jean-Pierre Aubin
Ivar Ekeland

February 1984
WP-84-5

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

PREFACE

This survey of nonsmooth analysis sets out to prove an inverse function theorem for set-valued maps. The inverse function theorem for the more usual smooth maps plays a very important role in the solution of many problems in pure and applied analysis, and we can expect such an adaptation of this theorem also to be of great value. For example, it can be used to solve convex minimization problems and to prove the Lipschitz behavior of its solutions when the natural parameters vary--a very important problem in marginal theory in economics.

ANDRZEJ WIERZBICKI
Chairman
System and Decision Sciences

CONTENTS

NONSMOOTH ANALYSIS

Jean-Pierre Aubin and Ivar Ekeland

INTRODUCTION

Nonlinear analysis must provide sufficient conditions for solving inclusions

$$(*) \qquad y \in F(x)$$

when F is a set-valued map from a Banach space X to a Banach space Y.

Our principal objective in this survey paper is to prove an Inverse Function Theorem for set-valued maps allowing us to say that when x_0 is a solution to

$$y_0 \in F(x_0) \quad ,$$

then there exist neighborhoods U of x_0 and V of y_0 such that inclusion (*) has solutions in U whenever y ranges over V. Furthermore, as in the smooth case, we require that the set of solutions $F^{-1}(y) \cap U$ of (*) depends in a Lipschitz manner on the data y. Since the Inverse Function Theorem for usual smooth maps plays such an important role in solving many problems of pure and applied analysis, we can expect an adaptation of the Inverse

Function Theorem to be very useful, more than just a generalization that had to be made. Actually, it can be used for solving convex minimization problems and proving the Lipschitz behavior of its solutions when the natural parameters vary. Economists claim that this problem is of utmost importance in their field (marginal theory). We shall take our inspiration from the smooth case where the sufficient condition is very simply stated: *the derivative at* x_0 *must be surjective.* The question arises: Can we define derivatives of set-valued maps such that the surjectivity of the derivative at (x_0,y_0) is sufficient for solving the surjectivity of F around y_0?

The answer to this question is one purpose for this paper. We now explain how we shall proceed to define derivatives of set-valued maps. We adopt the very first strategy, apparently suggested by Fermat, which defines the graph of the derivative to a smooth function as the tangent to the graph of this function. Therefore, we postpone questions about derivatives until after having tackled the matter of tangent spaces to subsets K of a Banach space X. They don't exist when K is no longer a smooth manifold. However, it is known in convex analysis that we can define in a natural way "tangent cones" to convex sets, which retain enough properties of tangent spaces to be quite useful. This is not enough, because most of the set-valued maps we shall meet have nonconvex graphs.

When K is neither smooth nor convex, there are many ways of defining "tangent cones", each one being as "natural" as the other. We shall retain only two concepts among the many candidates: the contingent cone and the tangent cone. Namely, they are defined in the following way: let x_0 belong to K.

The contingent cone, defined by

$$T_K(x_0) := \bigcap_{\varepsilon>0} \bigcap_{\alpha>0} \bigcup_{h\in]0,\alpha]} \left(\frac{1}{h}(K-x_0) + \varepsilon B\right)$$

was introduced by Bouligand in the early thirties, and the tangent cone defined by

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

was introduced by F.H. Clarke in 1975.

We see at once that the tangent cone $C_K(x)$ is contained in the contingent cone $T_K(x)$. They are both closed, and the tangent cone *is always convex*. We can say that they form a kind of "dipole", in the sense that the tangent cone $C_K(x_0)$ is the Kuratowski liminf of the contingent cones $T_K(x)$ when $x \to x_0$ (when the space X is finite dimensional). So, several properties of the tangent cone $C_K(x_0)$ at x_0 "diffuse" to generally weaker properties of the contingent cones $T_K(x)$ in a neighborhood of x_0. This "dipole" collapses to the usual tangent cone of K or the tangent space of K when K is convex and a smooth manifold respectively. We shall see in the sixth section that the contingent and tangent cones enjoy "dual properties".

These are some of the reasons for studying both contingent and tangent cones, and treating them as a pair rather than individuals.

Let F be a set-valued map from X to Y and (x_0,y_0) belong to its graph.

We define the *contingent derivative* $DF(x_0,y_0)$ as the *closed process* from X to Y whose graph is the contingent cone to the graph of F:

$$v_0 \in DF(x_0,y_0)(u_0) \iff (u_0,v_0) \in T_{\mathrm{graph}(F)}(x_0,y_0)$$

and the *derivative* $CF(x_0,y_0)$ as the *closed convex process* from X to Y whose graph is the tangent cone to the graph of F:

$$v_0 \in CF(x_0,y_0)(u_0) \iff (u_0,v_0) \in C_{\mathrm{graph}(F)}(x_0,y_0) \quad .$$

They may be different. Consider for instance the Lipschitz single-valued map π from $\mathbb{R}$ to $\mathbb{R}$ defined by

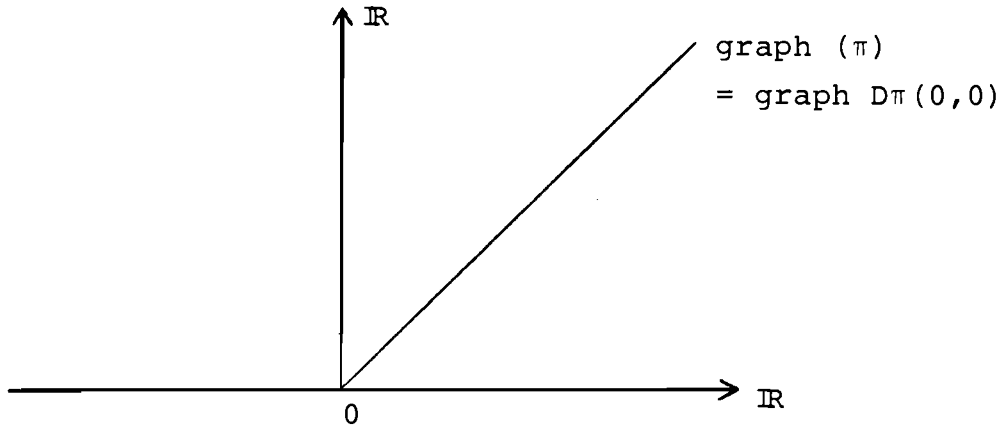
$$\pi(x) = 0 \text{ when } x \leq 0 \quad , \quad \pi(x) = x \text{ when } x \geq 0 \quad .$$

Then the contingent derivative of π at $(0,0)$ is defined by

$$D\pi(0,0)(u) = 0 \quad \text{when} \quad u \leq 0, \quad D\pi(0,0)(u) = u \quad \text{when} \quad u \geq 0$$

and the derivative of π at $(0,0)$ is defined by

$$C\pi(0,0)(u) = \emptyset \quad \text{when} \quad u \neq 0, \quad C\pi(0,0)(0) = 0 \quad .$$



This example shows that the price to pay for having a closed convex process as a derivative is sometimes too high.

These definitions provide intrinsic definitions of derivatives of single-valued maps defined on subsets K that may have an empty interior, as well as formulas for computing them when they are restrictions to K of a smooth map: when F is continuously differentiable on an open neighborhood of K, then the contingent derivatives and derivatives of the restriction $F|_K$ of F to K are the restrictions of the Jacobian ∇F of F to the contingent and tangent cones respectively:

i) $\quad D(F|_K)(x_0,F(x_0)) = \nabla F(x_0)|_{T_K(x_0)}$

ii) $\quad C(F|_K)(x_0,F(x_0)) = \nabla F(x_0)|_{C_K(x_0)}$

Also, these concepts of derivatives allow to compute the inverse of the derivative of a map, in particular, the inverse of the Jacobian of a single-valued map, because we infer immediately from the definitions that

$$\text{i)}\quad DF(x_0,y_0)^{-1} = D(F^{-1})(y_0,x_0)$$

$$\text{ii)}\quad CF(x_0,y_0)^{-1} = C(F^{-1})(y_0,x_0) \quad .$$

Since the derivative is a closed convex process, it is useful to distinguish its transpose $CF(x_0,y_0)^*$, a closed convex process from Y^* to X^*: we shall call it the *codifferential* of F at (x_0,y_0).

For real-valued functions, we can take into account the order relation, which is used in such problems as optimization problems (or Lyapunov functions, i.e., functions that decrease along the tranjectories of a dynamical system).

We associate with a proper function V from X to $\mathbb{R}\cup\{+\infty\}$ the set-valued map $\underset{\sim}{V}_+$ defined by $\underset{\sim}{V}_+(x) = V(x) + \mathbb{R}_+$ when $V(x) < +\infty$, $\underset{\sim}{V}_+(x) = \emptyset$ when $V(x) = +\infty$. We observe that there are numbers $D_+V(x_0)(u_0)$ and $C_+V(x_0)(u_0)$ such that

$$\text{i)}\quad D\underset{\sim}{V}_+(x_0,V(x_0))(u_0) = D_+V(x_0)(u_0) + \mathbb{R}_+$$

$$\text{ii)}\quad C\underset{\sim}{V}_+(x_0,V(x_0))(u_0) = C_+V(x_0)(u_0) + \mathbb{R}_+$$

where

$$-\infty \le D_+V(x_0)(u_0) \le C_+V(x_0)(u_0) \le +\infty \quad .$$

We shall say that the functions $D_+V(x_0)(\cdot)$ and $C_+V(x_0)(\cdot)$ from X to $\{-\infty\}\cup\mathbb{R}\cup\{+\infty\}$ are the epi-contingent derivative and epi-derivative of the function V.

In other words, the epigraphs of $D_+V(x_0)(\cdot)$ and $C_+V(x_0)(\cdot)$ are the contingent and tangent cones at the epigraph of V at $(x_0, V(x_0))$.

When the derivative $C_+V(x_0)(\cdot)$ is a proper function from X to $\mathbb{R}\cup\{+\infty\}$, it is convex, positively homogeneous and lower semi-continuous. This is, then, the support function of the closed convex subset

$$\partial V(x_0) := \{p \in X^* \mid \forall u \in X\,,\ \langle p,u\rangle \le C_+V(x_0)(u)\} \quad .$$

We shall call this subset the *generalized gradient* introduced by F.H. Clarke in 1975. Indeed, the terminology is justified by the fact that when V is continuously differentiable at x_0, then $\partial V(x_0) = \{\nabla V(x_0)\}$. We observe also that when V is convex, the generalized gradient coincides with the subdifferential $\partial V(x_0)$ of convex analysis.

It is then natural to consider the derivatives of the set-valued map $x \to \partial V(x)$ as candidates for the role of "second derivatives". Let p_0 belong to $\partial V(x_0)$; the derivative

$$C\partial V(x_0,p_0) := \partial^2 V(x_0,p_0)$$

is a closed convex process from X to X^*, which is monotone when V is convex.

These tangent cones and derivatives enjoy enough properties to make a decent calculus. But the main justification for including this study here is their use in the Inverse Function Theorem.

When X and Y are finite dimensional, it has a very simple formulation:

Let F *be a set-valued map with a closed graph and* (x_0,y_0) *belong to the graph of* F. *Assume that*

the derivative $CF(x_0,y_0)$ *of* F *at* (x_0,y_0) *is surjective.*

Then F^{-1} *is "pseudo-Lipschitz" around* (x_0,y_0) *in the sense that there exist a neighborhood* W *of* y_0, *two neighborhoods* U *and* V *of* x_0, $U \subset V$, *and a constant* $\ell > 0$ *such that*

i) $\forall y \in W,\ F^{-1}(y) \cap U \neq \emptyset$

ii) $\forall y_1, y_2 \in W,\ \delta(F^{-1}(y_1) \cap U, F^{-1}(y_2) \cap V) \leq \ell \| y_1 - y_2 \|$

where $\delta(A,B) := \sup_{x \in A} \inf_{y \in B} d(x,y)$.

It is itself a consequence of a more general Inverse Function Theorem, valid in infinite dimensional spaces, and involving surjectivity properties of the contingent derivative of F, not only at (x_0, y_0), but at all neighboring points.

We conclude this paper with a section devoted to the calculus of tangent cones, derivatives of set-valued maps and epi-derivatives of real-valued functions.

1. CONTINGENT AND TANGENT CONES

Let K be a nonempty subset of a Banach space X. We denote by εB and $\varepsilon \mathring{B}$ the ball (respectively, 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 *"contingent cone"* to K at x. ▲

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

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

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

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

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

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

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

For all $x \in X$, we have $T_X(x) = X$. We set $T_\emptyset(x) :+ \emptyset$. It is convenient to introduce the definition of the "liminf" of a family of subsets F(u).

Definition 2

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

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

▲

We observe that when the images of F are closed,

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

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

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

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

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

Definition 3

We say that the subset

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

is the tangent cone to K at x_0.

▲

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) for all sequences of elements $x_n \in X$, $h_n > 0$ converging to x_0 and 0, there exists a sequence of elements $u_n \in X$ converging to v such that $x_n + h_n u_n$ belongs to K for all n.

It is also characterized in the following way

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

We observe that when $x \in \text{Int}(K)$, then $C_K(x) = X$. For all $x \in X$, we have $C_X(x) = X$. We shall set $C_\emptyset(x) := \emptyset$. Tangent cones enjoy a very attractive property.

Proposition 4

The tangent cone $C_K(x_0)$ to K at x_0 is closed and convex.

▲

Proof.

Let v^1 and v^2 belong to $C_K(x_0)$. We take any sequence of elements $(x_n, h_n) \in K \times]0,\infty[$ converging to $(x_0, 0)$. There exists a sequence of elements v_n^1 converging to v^1 such that the elements $y_n := x_n + h_n v_n^1$ belong to K for all n. Since y_n converges to x_0, there exists a sequence of elements v_n^2 converging to v^2 such that $y_n + h_n v_n^2 = x_n + h_n(v_n^1 + v_n^2)$ belongs to K for all n. Since $v_n^1 + v_n^2$ converges to $v^1 + v^2$, we deduce that $v^1 + v^2$ belongs to $C_K(x_0)$. Hence the tangent cone is convex.

■

We note that

$$C_K(x_0) \subset T_K(x_0) \subset c\ell\left(\bigcup_{h>0} \frac{1}{h}(K - x_0)\right) \quad .$$

Proposition 5

If K is a convex subset, these three cones coincide:

$$C_K(x_0) = T_K(x_0) = c\ell \bigcup_{h>0} \frac{1}{h}(K - x_0) \quad . \tag{14}$$

▲

Proof

We have to prove that any $u_0 \in c\ell\left(\bigcup_{h>0} \frac{1}{h}(K - x_0)\right)$ belongs to $C_K(x_0)$. Let $\varepsilon > 0$ be fixed: there exist $y \in K$ and $\beta > 0$ such that $u_0 - \frac{1}{\beta}(y - x_0) \in \frac{\varepsilon}{2} B$. Let us take $\alpha := \beta\varepsilon/2$, x in $B_K(x_0, \alpha)$ and $h \in]0,\beta]$. We set $u := \frac{y-x}{\beta}$. Then $x + hu = (1 - \frac{h}{\beta})x + \frac{h}{\beta} y$ belongs to K because both x and y belong to K and $\frac{h}{\beta} \leq 1$. Also, $\|u - u_0\| \leq \frac{\|x - x_0\|}{\beta} + \|u_0 - \frac{y - x_0}{\beta}\| \leq \frac{\alpha}{\beta} + \frac{\varepsilon}{2} = \varepsilon$. Hence u_0 belongs to $C_K(x_0)$. ∎

These two cones may be different. Consider, for instance, the set K from $\mathbb{R}^2$, which is the graph of the map π from $\mathbb{R}$ defined by

$$\pi(x) = 0 \quad \text{when} \quad x \leq 0 \quad , \qquad \pi(x) = x \quad \text{when} \quad x \geq 0 \quad .$$

Then,

$$\text{if } x < 0, \quad C_K(x,0) = T_K(x,0) = \mathbb{R} \times \{0\}$$

$$\text{if } x = 0, \quad C_K(0,0) = \{0,0\}, \quad T_K(0,0) = (-\mathbb{R}_+ \times \{0\}) \cup \{u,u\}_{u \in \mathbb{R}_+}$$

$$\text{if } x > 0, \quad C_K(x,x) = T_K(x,x) = \{u,u\}_{u \in \mathbb{R}} \quad .$$

The tangent cone to K at (0,0) is convex, but trivial, whereas the contingent cone to K at (0,0) is nonconvex, but quite large.

We observe also that when K is a smooth manifold (of class C^1), then both the tangent cone and the contingent cone coincide with the usual tangent vector space to K at x of differential geometry.

The contingent and tangent cones are related by the following interesting relation.

Proposition 6

Assume that X is finite-dimensional. Then

$$(15) \qquad \forall x_0 \in K \quad , \qquad C_K(x_0) \subset \liminf_{x \to_K x_0} T_K(x) \quad .$$

▲

Proof

By definition of the tangent cone, we have

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

Let ε and α be fixed. It is clear that

$$\bigcup_{\beta>0} \bigcap_{x\in B_K(x_0,\alpha)} \bigcap_{h\in]0,\beta]} \left(\frac{1}{h}(K-x) + \varepsilon B\right)$$

$$\subset \bigcap_{x\in B_K(x_0,\alpha)} \bigcup_{\beta>0} \bigcap_{h\in]0,\beta]} \left(\frac{1}{h}(K-x) + \varepsilon B\right)$$

Since X is finite dimensional, we observe that any v in

$$\bigcup_{\beta>0} \bigcap_{h\in]0,\beta]} \left(\frac{1}{h}(K-x) + \varepsilon B\right) \text{ belongs to } T_K(x) + \varepsilon B \quad .$$

Indeed, there exist β and elements x_h such that

$$v \in \frac{x_h - x}{h} + \varepsilon B \quad \text{for} \quad h \le \beta.$$

A subsequence of $\frac{x_h - x}{h}$ converges to some w in $T_K(x)$. Hence

$$C_K(x_0) \subset \bigcap_{\varepsilon>0} \bigcup_{\alpha>0} \bigcap_{x\in B_K(x_0,\alpha)} (T_K(x) + \varepsilon B) = \liminf_{x \to_K x_0} T_K(x) .$$

■

The above inclusion is actually an equality.

Theorem 7. Let K be a nonempty weakly closed subset of a Hilbert space. The following inclusions hold true

$$(16) \qquad \liminf_{x \to_K x_0} T_K(x) \subset \liminf_{x \to_K x_0} (\overline{co}\, T_K(x)) \subset C_K(x_0) \quad .$$

When X is finite-dimensional, equalities hold true. Then 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 tangent cone to K at x_0.

■

The proof follows from the following lemmas.

Lemma 8. Let $K \subset X$ be a weakly closed subset. We denote by $\pi_K(y)$ the nonempty subset of elements $x \in K$ such that $\|x-y\| = d_K(y)$. We obtain the following

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

▲

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

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

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

Lemma 9. For any $y \in X$, we have

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

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

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

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

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

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

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

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

Lemma 9 ensues by taking the infimum when w ranges over $\overline{co}\ T_K(x)$ and x over $\pi_K(y)$.

Lemma 10. Let us consider the Lipschitz function f defined by $f(t) := \frac{1}{2} d_K(x+tv)^2$. For almost all $t \geq 0$, we have

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

▲

Proof of Theorem 7. 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 $2\alpha + \beta \|v_0\| \leq \eta$. 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 Lemma 10 that

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

because

$$d_K(x+tv_0) \leq t\|v_0\| \quad .$$

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

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

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 tangent cone $C_K(x_0)$. Then, by formula (15), we obtain:

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

When x is finite-dimensional, Proposition 6 implies that these three cones are equal.

■

The tangent cone $C_K(x_0)$ being a closed convex cone, it is equal to $C_K(x_0)^{--}$, its negative bipolar cone. This duality relation being quite useful, we introduce the following definition.

Definition 11

We shall say that the negative polar cone

(20) $$N_K(x_0) := C_K(x_0)^-$$

to the tangent cone to K at x_0 is the *normal cone to* K at x_0. ▲

2. CONTINGENT DERIVATIVES AND DERIVATIVES 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 set-valued map from* X *to* Y *and* (x_0, y_0) *belong belong to* ***graph*** (F).

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

In other words,

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

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

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

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 $(x_0, y_0) \in \text{graph}(F)$. ▲

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

We now 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:

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

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

When F is a single valued map, we set

(4) $$DF(x_0) := DF(x_0, F(x_0))$$

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

If F is C^1, then $DF(x_0)(u_0) = \nabla F(x_0)u_0$. When the graph of F is convex, we observe that u_0 belongs to $DF(x_0,y_0)(u_0)$ if and only if

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

<u>Proposition 2</u>

Assume that F is Lipschitz on a neighborhood of x_0 (belonging to Int Dom F). 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) $$\text{Dom } DF(x_0,y_0) = X \quad .$$

▲

Proof

a) The first statement follows from the fact that

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

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

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

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

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

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

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

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

Proposition 3

Let F be a single-valued map from an open subset Ω of X to Y of class C^1 and K be a nonempty subset of Ω containing x_0. Then

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

Proof

If F is a C^1 single-valued map at x_0 and u_0 belongs to $T_K(x_0)$, 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 $F|_K(x_0+h_n u_n) = F(x_0+h_n u_n) = F(x_0) + h_n\left(\nabla F(x_0)u_n + 0(h_n)\right)$

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 $(F|_K(x_0+h_nu_n) - F|_K(x_0))/h_n$. Therefore, $DF|_K(x_0,F(x_0))(u_0) = \nabla F(x_0)u_0$. ∎

We follow the same procedure in defining the 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 tangent cone $C_{graph(F)}(x_0,y_0)$ to the graph of F at (x_0,y_0). Briefly:

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

<u>Definition 4</u>

We shall say that the closed convex process $CF(x_0,y_0)$ from X to Y is the *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

$$(14)\quad \begin{cases} \forall \varepsilon_1,\varepsilon_2 > 0,\ \exists \alpha,\beta > 0 \text{ such that } \forall (x,y) \in B_{graph(F)}(x_0,y_0,\alpha)\ , \\ \forall h \in]0,\beta],\ \exists u \in u_0 + \varepsilon_1 B,\ v \in v_0 + \varepsilon_2 B \text{ such that } v \in \dfrac{F(x+hu) - y}{h} \end{cases}$$

or, equivalently, if and only if

$$(15)\quad \begin{cases} \text{for all sequences of elements } (x_n,y_n,h_n) \in graph(F) \times]0,\infty[\\ \text{converging to } (x_0,y_0,0), \text{ there exist sequences of elements} \\ u_n \text{ converging to } u_0 \text{ and } v_n \text{ converging to } v_0 \text{ such that} \\ y_n + h_nv_n \in F(x_n + h_nu_n) \text{ for all } n > 0\ . \end{cases}$$

The analytical formula involving "differential quotients" is quite complicated. It is simpler when F is locally Lipschitz: we begin with it.

Proposition 5

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

$$(16) \qquad \lim_{\substack{x \to_K x_0, h \to 0+}} d\left(v_0 , \frac{F(x+hu_0) - y_0}{h}\right) = 0 \quad .$$

▲

Remark

We observe that the domain of the derivative of a Lipschitz function is not necessarily the whole space, while the domain of the contingent derivative is the whole space when the dimension of Y is finite. Take for instance the map π associating to $x \in \mathbb{R}$, $\pi(x) := 0$ if $x \leq 0$ and $\pi(x) = x$ if $x > 0$. We saw that $C\pi(0,0)(u) = \emptyset$ when $u \neq 0$ and $C\pi(0,0)(0) = 0$, whereas $D\pi(0,0)(u) = \pi(u)$ for all $u \in \mathbb{R}$.

■

For the analytical formula in the general case, we need the following definition:

Definition 6

Let U and V be metric spaces and ϕ be a function from $U \times V$ to $\overline{\mathbb{R}}$. We set

$$(17) \qquad \limsup_{u \to u_0} \inf_{v \to v_0} \phi(u,v) := \sup_{\varepsilon > 0} \inf_{\eta > 0} \sup_{u \in B(u_0,\eta)} \inf_{v \in B(v_0,\varepsilon)} \phi(u,v) .$$

▲

Proposition 7

Let F be a proper set-valued map from X to Y and (x_0,y_0) belong to graph (F). Then v_0 belongs to the derivative $CF(x_0,y_0)(u_0)$ if and only if

$$(18) \qquad \left\{ \limsup_{\substack{(x,y) \to (x_0,y_0) \\ \text{graph } (F) \\ h \to 0+}} \inf_{u \to u_0} d\left(v_0 , \frac{F(x+hu) - y}{h}\right) = 0 \right.$$

▲

Proof of Propositions 5 and 7

Formula (14) can be written

$$\sup_{\varepsilon_1>0} \ \inf_{\alpha,\beta>0} \ \sup_{\substack{(x,y)\in B_{\mathrm{graph}(F)}(x_0,y_0;\alpha)\\ h\in]0,\beta]}} \ \inf_{u\in u_0+\varepsilon_1 B} d\left(v_0\,,\,\frac{F(x+hu)-y}{h}\right) = 0$$

This proves Proposition 5. When F is Lipschitz around x_0,

$$\left|\inf_{u\in u_0+\varepsilon_1 B} d\left(v_0,\frac{F(x+hu)-y}{h}\right) - d\left(v_0,\frac{F(x+hu_0)-y}{h}\right)\right| \leq \ell\varepsilon_1$$

and the above formulas become

$$\inf_{\alpha,\beta>0} \ \sup_{\substack{(x,y)\in B_{\mathrm{graph}(F)}(x_0,y_0,\alpha)\\ h\in]0,\beta]}} d\left(v_0,\frac{F(x+hu_0)-y}{h}\right) = 0 \quad .$$

■

When F is single-valued, we shall set

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

If F is continuously differentiable at x_0, we have

(20) $\quad CF(x_0) = \nabla F(x_0) \quad .$

Naturally, the formula for derivatives of inverses is obvious:

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

Proposition 8

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 containing x_0. Then

$$(22) \qquad CF|_K(x_0)u_0 = \begin{cases} \nabla F(x_0)u_0 & \text{if } u_0 \in C_K(x_0) \\ \emptyset & \text{if } u_0 \notin C_K(x_0) \end{cases}$$

▲

Proof

Let $(x_n,h_n) \in K \times]0,\infty[$ converge to $(x_0,0)$ in $K \times \mathring{\mathbb{R}}_+$. 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

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

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 $F|_K(x_n) + h_nv_n = F|_K(x+h_nu_n)$ for all n.

■

Since the derivative $CF(x_0,y_0)$ is a closed convex process, it is equal to its bi-transpose $CF(x_0,y_0)^{**}$. This suggests that we introduce the following definition.

Definition 9

We shall say that the transpose $CF(x_0,y_0)^*$ of the derivative of F at $(x_0,y_0) \in \text{graph}(F)$ is the codifferential of F at (x_0,y_0).

▲

It is a closed convex process from Y^* to X^* defined by

$$(23) \qquad \begin{cases} p_0 \in CF(x_0,y_0)^*(q_0) \text{ if and only if } \forall u \in X \quad , \\ \forall v \in CF(x_0,y_0)(u) \quad , \quad \langle p_0,u\rangle - \langle q_0,v\rangle \leq 0 \quad . \end{cases}$$

■

We mention an example of derivatives of a set-valued map that we shall use later.

Proposition 10

Let X and Y be Banach spaces, A be a continuously differentiable operator from an open subset Ω of X to Y and $L \subset \Omega$, $M \subset Y$ be closed subsets of X and Y respectively. Let F be the set-valued map from X to Y defined by

$$(24) \qquad F(x) := \begin{cases} A(x) - M & \text{when } x \in L \\ \emptyset & \text{when } x \notin L \end{cases}$$

Let (x_0, y_0) belong to the graph of F. 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 $v_0 \in \nabla A(x_0)u_0 - C_M(Ax_0 - y_0)$

▲

Proof

a) Let us prove that a) implies b). We take sequences $(x_n, z_n, h_n) \in L \times M \times]0,\infty[$ 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 $x_n + h_n u_n \in L$ and $A(x_n + h_n u_n) \in M + y_n + h_n v_n$ 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 $w_n := A(x_n + h_n u_n) - A(x_n) - v_n$ converges to $\nabla A(x_0)u_0 - v_0$ and because $z_n + h_n w_n$ belongs to M for all n.

b) Conversely, let us show that a) follows from b). We take a sequence $(x_n, y_n, h_n) \in \text{graph}(F) \times]0,\infty[$ converging to $(x_0, y_0, 0)$. There exists a sequence u_n converging to u_0 such that $x_n + h_n u_n$ 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 $Ax_n - y_n + h_n w_n \in M$ for all n. Then the sequence of elements $v_n := A(x_n+h_n u_n) - Ax_n - w_n$ converges to v_0 and satisfies $y_n + h_n v_n \in F(x_n+h_n u_n)$ for all n. ∎

Proposition 11

Let K be a 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) $u_0 \in C\ _K(x_0+p_0)(u_0+q_0)$.

The same result holds when the 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 $(y_n,h_n) \in X \times]0,\infty[$ 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 $(x_n,p_n,h_n) \in \text{graph } N_K \times]0,\infty[$ 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 12

Let us consider the set-valued map associating to x $\mathbb{R}^n_+$ 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

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

▲

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 $C\pi(0)(u) = \emptyset$ when $u \neq 0$ and $C\pi(0)(0) = 0$, we obtain the above corollary.

■

3. EPI-CONTINGENT DERIVATIVES AND EPI-DERIVATIVES OF REAL VALUED FUNCTIONS

We can use the concept of contingent derivatives and derivatives for single-valued maps V from DOM $V \subset X$ to $\mathbb{R}$: we obtain, for instance

$$(1) \qquad v_0 \in DV(x)(u_0) \iff \liminf_{\substack{h \to 0+ \\ u \to u_0}} \left| \frac{V(x+hu) - V(x)}{h} \right| = 0 .$$

In many problems such as minimization problems, the order relation plays an important role. This is the reason why we associate with a proper function $V : X \to \mathbb{R} \cup \{+\infty\}$ the set-valued map $\underset{\sim}{V}_+$ defined by $\underset{\sim}{V}_+(x) = V(x) + \mathbb{R}_+$ when $V(x) < +\infty$ and $\underset{\sim}{V}_+(x) = \emptyset$ when $V(x) = +\infty$. Its domain is the domain of V and its graph is the eopigraph of V. We consider its contingent derivative $D\underset{\sim}{V}_+(x,V(x))$, whose images are closed half-lines. Therefore, for all $u_0 \in X$, $D\underset{\sim}{V}_+(x,V(x)(u_0)$ is either $\mathbb{R}$, or a half line $[v_0,\infty[$, or empty. We set

$$(2) \qquad D_+V(x)(u) := \inf \{v \mid v \in D\underset{\sim}{V}_+(x,V(x))(u)\} .$$

It is equal to $-\infty$ if $D\underset{\sim}{V}_+(x,V(x)) = \mathbb{R}$, to v_0 if $D\underset{\sim}{V}_+(x,V(x))(u) = [v_0,\infty[$ and to $+\infty$ if $D\underset{\sim}{V}_+(x,V(x))(u) = \emptyset$.

<u>Definition 1</u>

We shall say that $D_+V(x)(u)$ is the "<u>epi-contingent derivative</u>" of V at x in the direction u.

We begin by computing epi-contingent derivatives.

▲

<u>Proposition 2</u>

If V is a proper function from X to $\mathbb{R} \cup \{+\infty\}$, then

$$(3) \qquad D_+V(x_0)(u_0) = \liminf_{\substack{h \to 0+ \\ u \to u_0}} \frac{V(x_0+hu) - V(x_0)}{h} .$$

The function $u \to D_+V(x_0)(u)$ is positively homogeneous and lower semi-continuous when $D_+V(x_0)(u) > -\infty$ for all $u \in X$.

▲

Proof

Indeed, let $v_0 \in D\underset{\sim}{V}_+(x_0, V(x_0))(u_0)$; then, $\forall \varepsilon_1 > 0$, $\varepsilon_2 > 0$, $\forall \alpha > 0$, there exist $u \in u_0 + \varepsilon_2 B$ and $h < \alpha$ such that $v_0 \in \dfrac{\underset{\sim}{V}_+(x_0+hu) - V(x_0)}{h} + \varepsilon_1 B$. This implies that

$$v_0 \geq \frac{V(x_0+hu) - V(x_0)}{h} - \varepsilon_1 \geq \inf_{h \leq \alpha} \inf_{\|u-u_0\| \leq \varepsilon_2} \frac{V(x_0+hu) - V(x_0)}{h} - \varepsilon_1 .$$

Therefore $v_0 \geq \liminf_{\substack{h \to 0+ \\ u \to u_0}} \dfrac{V(x_0+hu) - V(x_0)}{h} - \varepsilon_1$. Let us set for the time $a := \liminf_{\substack{h \to 0+ \\ u \to u_0}} \dfrac{V(x_0+hu) - V(x_0)}{h}$.

So, we have proved that $a \leq D_+V(x_0)(u_0)$. On the other hand, we know that for any $M > a$, for all α, $\delta > 0$, there exist $h < \alpha$, and $u \in u_0 + \delta B$ such that

$$\frac{V(x_0+hu) - V(x_0)}{h} \leq M .$$

Hence $M \in \dfrac{\underset{\sim}{V}_+(x_0+hu) - V(x_0)}{h}$. This proves that $a \in D\underset{\sim}{V}_+(x_0, V(x_0))(u_0)$. Since it is smaller than all the other ones, we infer that $a = D_+V(x_0)(u_0)$.

■

If V is C^1 at x_0, then

(4) $$\forall u_0 \in X \quad , \quad D_+V(x_0)(u_0) = \langle \nabla V(x_0), u_0 \rangle \quad .$$

If V is convex, then

(5) $$\forall u_0 \in X \quad , \quad D_+V(x_0)(u_0) = \liminf_{u \to u_0} \left(\inf_{h>0} \frac{V(x_0+hu) - V(x_0)}{h} \right) .$$

We deduce from Propositions 2-2 and 2-3 the following statements.

Proposition 3

Let us assume that V is Lipschitz on a neighborhood of $x_0 \in \text{Int Dom } V$. Then

(6) $$\forall u_0 \in X \ , \quad D_+V(x_0)(u_0) = \liminf_{h \to 0+} \frac{V(x_0+hu_0) - V(x_0)}{h}$$

and the epi-contingent derivative is finite.

▲

Proposition 4

Let V be a proper function from X to $\mathbb{R} \cup \{+\infty\}$ and K be a subset of X. Let $V|_K$ denote the restriction of V to K (in the sense that $V|_K(x)$ equals $V(x)$ when $x \in K, \infty$ when $x \notin K$).

Then

(7) $$\forall x_0 \in K, \quad \forall v_0 \in T_K(x_0), \quad D_+V(x_0)(u_0) \leq D_+V|_K(x_0)(u_0) \quad .$$

If V is C^1 at x_0, we have

(8) $$D_+V|_K(x_0)(u_0) = \begin{cases} \langle \nabla V(x_0), u_0 \rangle & \text{if} \quad u_0 \in T_K(x_0) \\ +\infty & \text{if} \quad u_0 \notin T_K(x_0) \end{cases}$$

▲

We state the obvious property of the epi-contingent derivative at a minimizer.

Proposition 5

Let V be a proper function from a Banach space X to $\mathbb{R}\cup\{+\infty\}$. If $\bar{x}$ Dom V minimizes V on X, then

$$(9) \qquad \forall u \in X \quad , \qquad 0 \leq D_+V(\bar{x})(u) \quad .$$

▲

More generally, the ε-variational principle of Ekeland can take the following form:

Theorem 6

Let V be a proper lower semicontinuous function bounded below from a Banach space X to $\mathbb{R}\cup\{+\infty\}$ and x_0 belong to Dom V. Then, for any $\varepsilon > 0$, there exists $x_\varepsilon \in \text{Dom } V$ satisfying

$$(10) \qquad \begin{cases} \text{i)} & V(x_\varepsilon) + \varepsilon\|x_\varepsilon - x_0\| \leq V(x_0) \\ \text{ii)} & \forall u \in X, \quad 0 \leq D_+V(x_\varepsilon)(u) + \varepsilon\|u\| \end{cases}$$

▲

Proof

By Ekeland's theorem (the ε-variational principle), there exists $x_\varepsilon \in \text{Dom } V$ satisfying (10) i) and $V(x_\varepsilon) = \min_{x\in X}[V(x) + \|x - x_\varepsilon\|]$.

Let $u \in \text{Dom } D_+V(x_\varepsilon)$. Then, for any $\eta > 0$, $\delta > 0$, $\alpha > 0$, there exist $h \leq \alpha$ and $v \in u + \delta B$ such that

$$\frac{V(x_\varepsilon + hv) - V(x_\varepsilon)}{h} \leq D_+V(x_\varepsilon)(u) + \eta \quad .$$

Ekeland's theorem implies:

$$-\varepsilon\delta - \varepsilon\|u\| \leq -\varepsilon\|v\| \leq \frac{V(x_\varepsilon + hv) - V(x_\varepsilon)}{h} \quad .$$

Therefore, we infer that

$$0 \leq D_+V(x_\varepsilon)(u) + \varepsilon\|u\| + \varepsilon\delta + \eta \quad .$$

By letting δ and η converge to 0, we obtain the desired inequality. ■

We define in the same way epi-derivatives of functions V from X to $\mathbb{R}\cup\{+\infty\}$. Since the images of the derivative $C\underset{\sim}{V}_+(x_0,V(x_0))$ are either $\mathbb{R}$, or a half-line $[v_0,\infty[$, or empty, we set:

$$(11) \qquad C_+V(x_0)(u_0) := \inf\{v \mid v \in C\underset{\sim}{V}_+(x_0,V(x_0))(u_0)\} \quad .$$

It is equal to $-\infty$ when $C\underset{\sim}{V}_+(x_0,V(x_0)) = \mathbb{R}$, to v_0 when $C\underset{\sim}{V}_+(x_0,V(x_0))(u_0) = [v_0,\infty[$ and to $+\infty$ when $C\underset{\sim}{V}_+(x_0,V(x_0))(u_0) = \emptyset$.

<u>Definition 7</u>

We shall say that $C_+V(x_0)(u_0)$ is the "epi-derivative" of V at x_0 in the direction u_0. ▲

The epigraph of $u \to C_+V(x_0)(u)$ is a closed convex cone because it is the graph of the set-valued map $u \to CV_+(x_0,V(x_0))(u)$, which is a closed convex process. We deduce at once the following important property.

<u>Proposition 8</u>

The epi-derivative $u \to C_+V(x_0)(u)$ is a positively homogeneous lower semicontinuous convex function when $C_+V(x_0)(u) > -\infty$ for all $u \in X$. ▲

It is easy to check that the co-differential of $\underset{\sim}{V}_+$ at $(x_0,V(x_0))$ is a closed convex process from $\mathbb{R}$ to X^*, defined by its values $C\underset{\sim}{V}_+(x_0,V(x_0))^*(-1)$ and $C\underset{\sim}{V}_+(x_0,V(x_0))^*(1)$. We observe that $C\underset{\sim}{V}_+(x_0,V(x_0))^*(-1) = \emptyset$ and that the support function of $C\underset{\sim}{V}_+(x_0,V(x_0))^*(1)$ (when it is not empty), is equal to $C_+V(x_0)(\cdot)$. ■

<u>Definition 9</u>

We say that the closed convex subset of X^* defined by:

$$(12) \qquad \begin{cases} \partial V(x_0) := C\underset{\sim}{V}_+(x_0,V(x_0))^*(1) \\ \\ =\{p \in X^* \mid \forall u \in X, \langle p,u \rangle \leq C_+V(x_0)(u)\} \end{cases}$$

is the <u>generalized</u> <u>gradient</u> of V at x_0. ▲

It is empty whenever there exists a direction u_0 for which $C_+V(x_0)(u_0) = -\infty$.

When V is continuously differentiable at x_0, then $\forall u_0 \in X$, $C_+V(x_0)(u_0) = \langle \nabla V(x_0), u_0 \rangle$ and, consequently,

$$\partial V(x_0) = \{\nabla V(x_0)\} \quad .$$

This motivates the term "generalized gradient". When V is convex, it coincides with the subdifferential of V at x_0.

Proposition 10

Let us assume that V is Lipschitz on a neighborhood of $x_0 \in \text{Int Dom } V$. Then

$$(13) \qquad \forall u_0 \in X, \quad C_+V(x_0)(u_0) = \limsup_{\substack{x \to x_0 \\ h \to 0+}} \frac{V(x+hu_0) - V(x)}{h}$$

and the epi-derivative is finite. Furthermore, the following properties hold true

$$(14) \quad \begin{cases} \text{i)} & \forall u_0 \in X, \ (x,u) \to C_+V(x_0)(u) \text{ is upper semi-continuous at } (x_0,u) \\ \text{ii)} & u \to C_+V(x_0)(u) \text{ is continuous} \\ \text{iii)} & C_+(-V)(x_0)(u) = C_+V(x_0)(-u) \quad . \end{cases}$$

In terms of generalized gradients, these properties become

$$(15) \quad \begin{cases} \text{i)} & x \to \partial V(x) \text{ is upper hemicontinuous at } x_0 \\ \text{ii)} & \partial V(x_0) \text{ is (closed convex and) bounded} \\ \text{iii)} & \partial(-V)(x_0) = -\partial V(x_0) \quad . \end{cases}$$

▲

Proof

Since V is Lipschitz on a neighborhood of $x_0 \in \text{Int Dom } V$, there exist $\alpha_0 > 0$ and $\ell > 0$ such that, for any α, β, η satisfying

$\alpha + \beta(\|u_0\|+\eta) \leq \alpha_0$, we have: $\forall x \in x_0 + \alpha B$, $\forall h \in]0,\beta]$, $\forall u \in u_0 + \eta B$,

(16) $$\left|\frac{V(x+hu) - V(x)}{h}\right| \leq \ell(\|u_0\|+\eta)$$

a) Let v_0 belong to the derivative $C\underset{\sim}{V}_+(x_0,V(x_0))$ of V_+ at $(x_0,V(x_0))$. This means that for all $\varepsilon, \eta > 0$, there exist $\alpha, \beta > 0$ such that, for all $x \in x_0 + \alpha B$, $\forall h \in]0,\beta]$, there exists $u \in u_0 + \eta B$ such that

$$v_0 \in \frac{\underset{\sim}{V}_+(x+hu) - V(x)}{h} + \varepsilon B \quad .$$

Hence $v_0 \geq \dfrac{V(x+hu) - V(x)}{h} - \varepsilon \geq \dfrac{V(x+hu_0) - V(x)}{h} - \varepsilon - \ell\eta$ (because V is Lipschitz around x_0).

Consequently, $v_0 \geq \limsup\limits_{\substack{x \to x_0 \\ h \to 0+}} \dfrac{V(x+hu_0) - V(x)}{h}$ and thus

$$C_+V(x_0)(u_0) \geq \limsup_{\substack{x \to x_0 \\ h \to 0+}} \frac{V(x+hu_0) - V(x)}{h} \quad .$$

Conversely, let us set $a := \limsup\limits_{\substack{x \to x_0 \\ h \to 0+}} \dfrac{V(x+hu_0) - V(x)}{h}$,

which is finite by inequality (16). Then we can associate to any $\varepsilon > 0$ constants $\alpha, \beta > 0$ such that

$$a + \varepsilon \geq \frac{V(x+hu_0) - V(x)}{h} \quad .$$

This implies that a belongs to $CV_+(x_0,V(x_0))$. Hence formula (13) ensues.

b) The upper semicontinuity of $(x,u) \to C_+V(x)(u)$ at (x_0,u_0) follows at once from formula (13).

Also, inequality (16) implies that

$$(17) \qquad C_+V(x_0)(u) \leq \ell\|u\|$$

and thus, that $u \to C_+V(x_0)(u)$ is continuous. To prove (14) iii) we observe that

$$\frac{-V(x+hu_0) - (-V(x))}{h} = \frac{V((x+hu_0) + h(-u_0)) - V(x+hu_0)}{h} .$$

Since $x + hu_0$ is in a neighborhood of x_0 when x is a neighborhood of x_0 and h is small, we deduce that $C_+(-V)(x_0)(u_0) = C_+V(x_0)(-u_0)$.

c) Since $C_+V(x_0)(\cdot)$ is proper, it is the support function of $\partial V(x_0)$. ∎

Remark

More generally, we can prove the following formula for epi-derivatives of arbitrary functions.

For that purpose, it is expedient to use the notation:

$$(18) \qquad (x,\lambda) \downarrow x_0 \iff \lambda \geq V(x) \quad , \quad x \to x_0 \quad \text{and} \quad \lambda \to V(x_0)$$

and the definition 2.6 of lim sup inf.

Proposition 11

Let x_0 belong to the domain of a function V from X to $\mathbb{R} \cup \{+\infty\}$. Then

$$(19) \qquad C_+V(x_0)(u_0) = \limsup_{\substack{(x,\lambda)\downarrow x_0 \\ h \to 0+}} \inf_{u \to u_0} \frac{V(x+hu) - \lambda}{h} .$$

The proof is left as an exercise.

When V is lower semicontinuous at x_0, formula (19) becomes

$$(20) \qquad C_+V(x_0)(u_0) = \limsup_{\substack{x \to x_0 \\ V(x) \to V(x_0) \\ h \to 0+}} \inf_{u \to u_0} \frac{V(x+hu) - V(x)}{h} .$$

It may be useful to use another concept of derivative, easier to manipulate than the epi-derivative.

Definition 12

Let V be a proper function from X to $\mathbb{R} \cup \{+\infty\}$ and let x_0 belong to Dom V. We set

$$(21) \qquad B_+V(x_0)(u_0) := \limsup_{\substack{(x,\lambda) \downarrow x_0 \\ u \to u_0 \\ h \to 0+}} \frac{V(x+hu) - \lambda}{h} .$$

We shall say that $B_+V(x_0)(u_0)$ is the strict epi-derivative of V at x_0 in the direction of u_0 and that V is strictly epi-differentiable at x_0 if the function $u \to B_+V(x_0)(u)$ is a proper function from X to $\mathbb{R} \cup \{+\infty\}$. ▲

We always have

$$(22) \qquad \forall u \in X \quad , \quad D_+V(x_0)(u) \le C_+V(x_0)(u) \le B_+V(x_0)(u) .$$

Clearly, a function V which is Lipschitz around x_0 is strictly epi-differentiable at x_0. The introduction of this concept is justified by the following result.

Proposition 13

Let us assume that the function V is strictly epi-differentiable at $x_0 \in$ Dom V.

Then

(23) $$\text{Dom } B_+V(x_0) = \text{Int Dom } C_+V(x_0)$$

and

(24) $$\forall u_0 \in \text{Dom } C_+V(x_0), \quad C_+V(x_0)(u_0) = \liminf_{u \to u_0} B_+V(x_0)(u_0) .$$

Furthermore, for any $u_0 \in \text{Int Dom } C_+V(x_0)$,

(25) $$\begin{cases} \text{i)} & (x,u) \to C_+V(x)(u) \text{ is upper semicontinuous at } (x_0,u_0) \\ \text{ii)} & u \to C_+V(x_0)(u) \text{ is continuous at } u_0 \end{cases} .$$

If we assume that $\text{Dom } B_+V(x_0) = X$, then

(26) $$\partial(-V)(x_0) = -\partial V(x_0) .$$

▲

<u>Proof</u>

a) Let u_0 belong to the domain of $B_+V(x_0)$. Equation (21) implies at once that $\text{Dom } B_+V(x_0)$ is open and that $(x,u) \to B_+V(x)(u)$ is upper semicontinuous at (x_0,u_0).

b) Formula (21) implies that

(27) $$B_+V(x_0)(u_0+u_1) \leq B_+V(x_0)(u_0) + C_+V(x_0)(u_1) .$$

We deduce that any u interior to the domain of $C_+V(x_0)$ belongs to the domain of $B_+V(x_0)$. For that purpose, take $u_0 \in \text{Dom } B_+V(x_0)$ and $\lambda > 0$ such that $u - \lambda u_0$ belongs to the domain of $C_+V(x_0)$. Then inequality (27) implies that

$$B_+V(x_0)(u) \leq C_+V(x_0)(u-\lambda u_0) + \lambda B_+V(x_0)(u_0) < +\infty ,$$

i.e., that u belongs to the domain of $B_+V(x_0)$. Hence the domain of $B_+V(x_0)$ coincides with the interior of the domain of $C_+V(x_0)$. Inequality (27) implies also that the epigraph of $B_+V(x_0)$ is dense in the epigraph of $C_+V(x_0)$. Consequently,

$$\liminf_{u \to u_0} B_+V(x_0)(u) \leq C_+V(x_0)(u_0) \quad .$$

Since $u \to C_+V(x_0)(u)$ is lower semicontinuous, equality (24) ensues. Furthermore, by letting λ go to 0 in the above inequality, we get:

(28) $$\forall u \in \text{Dom } B_+V(x_0) \quad , \quad B_+V(x_0)(u) = C_+V(x_0)(u) \quad .$$

Inequality (24) implies that

(29) $$\partial V(x_0) = \{p \in X^* \mid \forall u \in X, \langle p,u \rangle \leq B_+V(x_0)(u)\} \quad .$$

Hence property (26) follows from

(30) $$\forall u \in \text{Dom } B_+V(x_0) \quad , \quad B_+V(x_0)(-u) = B_+(-V)(x_0)(u) \quad .$$

For proving it, let us set $v_0 := B_+V(x_0)(-u_0)$; for all $\varepsilon > 0$, there exist α_0, β_0, $\eta_0 > 0$ such that, for all $y \in \text{Dom } V \cap (x_0+\alpha_0 B)$, $h \in]0,\beta u]$, $u \in u_0 + \eta_0 B$, $\dfrac{V(y-hu) - V(y)}{h} \leq v_0 + \varepsilon$.

Let us take $\alpha \in]0,\alpha_0]$, $\beta \in]0,\beta_0]$ and $\eta \in]0,\beta_0]$ such that $\alpha + \beta(\|u_0\|+\eta) \leq \alpha_0$. Hence, for all $x \in \text{Dom } V \cap (x_0+\alpha B)$, $\lambda \in V(x_0) + \alpha B$, satisfying $\lambda \geq -V(x)$, $h \in]0,\beta]$, $u \in u_0 + \eta B$, we have, by setting $y := x + hu$,

$$\frac{-V(x+hu) - \lambda}{h} \leq \frac{V(x) - V(x+hu)}{h} = \frac{V(y-hu) - V(y)}{h} \leq v_0 + \varepsilon$$

because y belongs to $\text{Dom } V \cap (x_0+\alpha_0 B)$. This implies that $B_+(-V)(x_0) \leq v_0 := B_+V(x_0)(-u_0)$. By exchanging the roles of V and -V, we have proved equality (30).

∎

Proposition 14

Let P be a closed convex cone. Assume that V is nonincreasing with respect to P in the sense that $V(x+y) \leq V(x)$ for all $y \in P$. Then let x_0 belong to Dom V. Then

$$(31) \qquad \forall u_0 \in P \quad , \quad C_+V(x_0)(u_0) \leq 0 \quad .$$

and

$$(32) \qquad \partial V(x_0) \subset P^- \quad .$$

If Int $P \neq \emptyset$, then

$$(33) \qquad \forall u_0 \in \text{Int } P \quad , \quad B_+V(x_0)(u_0) \leq 0 \quad .$$

Proof

Indeed, for any $x \in x_0 + \alpha B$, $\lambda \in V(x_0) + \alpha B$, $\lambda \geq V(x)$, $h \in]0,\beta[$, $u_0 \in P$, we have

$$\frac{V(x+hu_0) - \lambda}{h} \leq \frac{V(x+hu_0) - V(x)}{h} \leq 0$$

because V is nonincreasing. Hence $C_+V(x_0)(u_0) \leq 0$. If u_0 belongs to the interior of P, there exists $\eta_0 > 0$ such that $u_0 + \eta_0 B \subset P$. Hence, for all $u \in u_0 + \eta_0 B$, we would have

$$\frac{V(x+hu) - \lambda}{h} \leq \frac{V(x+hu) - V(x)}{h} \leq 0$$

and thus, $B_+V(x_0)(u_0) \leq 0$. ∎

We deduced the property of epi-contingent derivatives and epi-derivatives from the properties of contingent cones and tangent cones. Conversely, we can derive properties of contingent cones and tangent cones from those of contingent derivatives and epi-derivatives because we remark that when x_0 belongs to a subset K,

(34) $$D_{+}\psi_K(x_0,0) = \psi_{T_K}(x_0) \quad \text{and} \quad C_{+}\psi_K(x_0,0) = \psi_{C_K}(x_0) \quad ,$$

where ψ_L denotes the indicator of the subset L. We also mention the following useful properties.

Proposition 15

a) If $p \in X^*$ satisfies $\langle p,x_0 \rangle = \max_{y\in K} \langle p,y \rangle$, then p belongs to the normal cone $N_K(x_0)$.

b) Assume now that X is a Hilbert space.

If $y \notin \overline{K}$ and if $x \in \pi_{\overline{K}}(y)$ is a projection of y to $\overline{K}$, then y-x belongs to the normal cone $N_K(x)$.

▲

Proof.

a) If $p \in X^*$ satisfies $\langle p,x_0 \rangle = \max_{y\in K} \langle p,y \rangle$, then $x_0 \in K$ minimizes on K the linear functional $x \to \langle p,x \rangle$ and thus, $0 \in \partial(-p|_K)(x_0) \subset -p + N_K(x_0)$ by propositions 4 and 5.

b) Since the function $V : x \to \|y-x\| =: V(x)$ is continuously differentiable at all $x \neq y$ and since $x \in \pi_{\overline{K}}(y)$ minimizes V on K, we deduce that $0 \in \partial(V|_K)(x) \subset \nabla V(x) + N_K(x) = \frac{x-y}{\|x-y\|} + N_K(x)$. Hence, $y-x \in N_K(x)$.

Proposition 16

Let V be a proper upper semicontinuous function from X to $\mathbb{R} \cup \{+\infty\}$. We set

(35) $$K := \{x \in X | V(x) \leq c\} \quad .$$

Let $x_0 \in K$ satisfy $V(x_0) = c$. Then

(36) $$T_K(x) \subset \{v \in X | D_{+}V(x)(v) \leq 0\} \quad .$$

If we assume that

(37) $\exists\, u_0 \in X$ such that $C_+V(x_0)(u_0) < 0$

and V is upper semi-continuous at x_0, then inclusion:

(38) $\{u \in X \mid C_+V(x_0)(u) \leq 0\} \subset C_K(x_0)$

holds true.

▲

Proof

We first check that if u_0 satisfies $C_+V(x_0)(u_0) < 0$, then u_0 belongs to $C_K(x_0)$. Let us set $v_0 := -C_+V(x_0)(u_0) > 0$. For all $\varepsilon \in]0, v_0[$, there exist $\alpha > 0$ and $\beta > 0$ such that, for all $x \in x_0 + \alpha B$, $h \in]0,\beta[$, there exists $u \in u_0 + \varepsilon B$ such that $V(x+hu) \leq V(x) + h(-v_0+\varepsilon)$. Hence, for all $x \in B_K(x_0,\alpha)$, $h \in]0,\beta[$, there exists $u \in u_0 + \varepsilon B$ such that $V(x+hu) \leq V(x_0)$, i.e., $x + hu \in K$. Therefore, u_0 belongs to $C_K(x_0)$.

Now, if u satisfies $C_+V(x_0)(u) \leq 0$, then, for all $\lambda \in]0,1[$, $u_\lambda := (1-\lambda)u + \lambda u_0$ satisfies $C_+V(x_0)(u_\lambda) < 0$ by convexity and thus $u_\lambda \in C_K(u_0)$. Hence we deduce that u belongs to $C_K(u_0)$ by letting λ converge to 0.

4. GENERALIZED SECOND DERIVATIVES OF REAL-VALUED FUNCTIONS

Let V be a proper function from X to $\mathbb{R}$ $\{+\infty\}$. We consider the set-valued map ∂V from X to X^* associating to each $x_0 \in X$ the generalized gradient of V at x_0.

Therefore, if (x_0,p_0) belongs to the graph of ∂V, the derivative $C(\partial V)(x_0,p_0)$ of ∂V at (x_0,p_0) plays the role of a second derivative of V.

Definition 1

We shall say that the derivative

(1) $$\partial^2 V(x_0,p_0) := C(\partial V)(x_0,p_0)$$

of the map ∂V at $(x_0,p_0) \in \text{graph}(\partial V)$ is the <u>generalized</u> <u>second</u> <u>derivative</u> of V at (x_0,p_0). ▲

Therefore, $\partial^2 V(x_0,p_0)$ is a <u>closed</u> <u>convex</u> <u>process</u> from X to X^*.

It is clear that when V is twice continuously differentiable at x_0, then $p_0 = \nabla V(x_0)$ and $\partial^2 V(x_0;p_0)$ coincides with the Hessian $\nabla^2 V(x_0)$, mapping X to X^*.

<u>Proposition 2</u>

Let V be a proper lower semicontinuous convex function from X to $\mathbb{R} \cup \{+\infty\}$ and V^* its conjugate function. Then $\partial^2 V(x_0,p_0)$ is a monotone closed convex process and

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

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

(3)
$$\text{i)} \quad D_+V(x_0)(u_0) = \langle p_0,u_0 \rangle$$
$$\text{ii)} \quad D_+V^*(p_0)(q_0) = \langle p_0,x_0 \rangle$$

<u>Proof</u>

a) Let (u^i,q^i) (i=1,2) be two pairs of the graph of $\partial^2 V(x_0,p_0)$. Let h_n converge to 0+. Then we know that there exist sequences of elements u_n^i and v_n^i converging to u^i and v^i such that $(x_0+h_n u_n^i, p_0+h_n q_n^i)$ belong to the graph of ∂V for i = 1,2. Since the graph of ∂V is monotone, we deduce that

$$h_n^2 \langle q_n^1-q_n^2, u_n^1-u_n^2 \rangle = \langle p_0+h_n q_n^1-(p_0+h_n q_n^2), x_0+h_n u_n^1-(x_0+h_n u_n^2) \rangle \geq 0$$

Hence $\partial^2 V(x_0,p_0)$ is monotone.

Inequality (2) is straightforward since ∂V^* is the inverse of ∂V.

5. THE INVERSE FUNCTION THEOREM

We denote by ρB and ρB^{o} the closed and open balls of radius ρ respectively. We set $\mathrm{d}(A,B) := \sup_{x\in A} \inf_{y\in B} \|x-y\|$. Note that $\mathrm{d}(A,B) = 0$ means that A is contained in the closure of B.

We shall extend the usual Inverse Function Theorem for continuously differentiable single-valued maps to the case of set-valued maps. We need the following definition.

Definition 1

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

(1)
i) $\forall x \in W, \quad F(x) \cap U \neq \emptyset$

ii) $\forall x_1,x_2 \in W, \ \mathrm{d}(F(x_1) \cap U, F(x_2) \cap V) \leq \ell \|x_1-x_2\|$.

▲

Note that, if F is single-valued on W, it will be pseudo-Lipschitz if and only if it is ℓ-Lipschitz on W.

Theorem 2

Let F be a proper set-valued map with closed graph from X to Y and let (x_0,y_0) belong to graph (F). We assume that

(2)
i) both X and Y are finite dimensional

ii) the derivative $CF(x_0,y_0)$ of F at (x_0,y_0) is surjective (i.e. $\mathrm{Im}\ CF(x_0,y_0) = Y$).

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

▲

We start with the following lemma.

Lemma 3

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

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

Then, for all $\alpha > 0$, there exist constants $c > 0$ 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

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

▲

Proof

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

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

Let us introduce the subset

(6) $K := (B \times \gamma B) \cap \text{graph}\ CF(x_0, y_0)$.

Since the spaces X and Y are finite-dimensional, the subset K is a compact subset of the tangent cone $C_F(x_0, y_0)$ to F, the graph of F, at (x_0, y_0), which is the lim inf of the contingent cones $T_F(x,y)$ to the graph of F at points (x,y) converging to (x_0, y_0). Hence we can associate to every $\alpha > 0$ a positive number η 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 (5), 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 $v \in DF(x,y)(u) + w$, $\|u\| \leq \frac{1+\alpha}{\gamma}\|v\|$ and $\|w\| \leq \frac{\alpha}{\gamma}\|v\|$. ∎

Theorem 2 then becomes a consequence of the following general Inverse Function Theorem, valid in all Banach spaces.

<u>Theorem 4</u>

Let F be a proper set-valued map with closed graph 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

$$(7) \quad \begin{cases} \text{i)} & v \in DF(x,y)(u) + w \\ \text{ii)} & \|u\| \leq c\|v\| \quad \text{and} \quad \|w\| \leq \alpha\|v\| \end{cases}$$

Let us set

$$(8) \quad \begin{aligned} & r := \frac{\eta(1-\alpha)}{3(1+\alpha+c)},\ F_0^{-1}(y) := F^{-1}(y) \cap \left(x_0 + \frac{c+2\alpha}{1-\alpha} rB\right) \\ & \text{and } F_1^{-1}(y) := F^{-1}(y) \cap \left(x_0 + \frac{3(c+2\alpha)}{1-\alpha} rB\right) \end{aligned}$$

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

$$(9) \quad \begin{cases} \text{i)} & \forall y \in y_0 + rB,\ F_0^{-1}(y) \neq \emptyset \\ \text{ii)} & \forall y_1,\ y_2 \in y_0 + r\mathring{B},\ d\left(F_0^{-1}(y_1), F_1^{-1}(y_2)\right) \\ & \leq \frac{c+2\alpha}{1-\alpha}\|y_1-y_2\| \end{cases}$$

▲

Proof

Let y_1 and y_2 belong to the open ball $y_0 + rB$. Assume for the time being that there exists x_1 satisfying

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

(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

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

We apply Ekeland's Theorem 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

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

Inequality (12)i) implies that

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

Therefore

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

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

Consequently, we can use property (7) with $v := y_2-\bar{y}$: there exist u and w satisfying

$$(13) \quad \begin{cases} \text{i)} & y_2 - \bar{y} \in DF(\bar{x},\bar{y})(u) + w \\ \text{ii)} & \|u\| \leq c\|y_2-\bar{y}\| \text{ and } \|w\| \leq \alpha\|y_2-\bar{y}\| \end{cases} .$$

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 (12)ii), we obtain

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

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

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

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

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 (12)i), we get

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

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

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

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

$$(14) \qquad d(x_1,F_1^{-1}(y_2)) \leq \ell\|y_1-y_2\| \quad .$$

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

(15) $$\forall y_2 \in y_0 + rB, \quad x_2 \in F_0^{-1}(y_2) := F^{-1}(y_2) \cap x_0 + \frac{c+2\alpha}{1-\alpha} rB)$$

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

In other words, the set-valued map F_0^{-1} has non empty images when y ranges over the open ball $y_0 + rB$. Inequality (14) implies that

$$\mathbb{d}(F_0^{-1}(y_1),F_1^{-1}(y_1)) := \sup_{x_1 \in F_0^{-1}(y_1)} d(x_1,F_1^{-1}(y_2)) \leq \frac{c+2\alpha}{1-\alpha}\|y_1-y_2\| \quad . \blacksquare$$

As a first consequence, we obtain the usual Liusternik Theorem.

Corollary 5

Let f be a continuously differentiable map from an open subset Ω of a Banach space X to a Banach space Y. Assume that for $x_0 \in \Omega$,

(16) $\nabla f(x_0)$ is surjective .

Then there exist neighborhoods U and V of x_0, $U \subset V$, and W of $f(x_0)$ such that, for all $y \in W$, there exists a solution $x \in U$ to the equation $f(x) = y$ and such that

(17) $$\forall y_1,y_2 \in W, \ \mathbb{d}(f^{-1}(y_1) \cap U, \ f^{-1}(y_2) \ V) \leq \ell\|y_1-y_2\| \quad .$$

Proof

Let K be a closed neighborhood of x_0 contained in Ω. We apply Theorem 4 to the restriction F of f to K. Since $\nabla f(x_0)$ is surjective, there exists a constant $c > 0$ such that for all $v \in Y$, there exists a solution u of the equation $\nabla f(x_0)u = v$ satisfying $\|u\| \leq c\|v\|$. Let $\alpha > 0$ be given and $\eta > 0$ such that $\|\nabla f(x)-\nabla f(x_0)\| \leq \alpha$ when $x \in B(x_0,\eta) \subset \mathrm{Int}\ K$. Then the assumptions of Theorem 4 are satisfied because $v = \nabla f(x)u + w$ where $\|u\| \leq c\|v\|$ and

$$w := (\nabla f(x) - \nabla f(x_0))v \quad \text{is such that} \quad \|w\| \leq \alpha\|v\| \quad . \quad \blacksquare$$

By taking $n = \infty$, we obtain the following corollary.

Corollary 6 (normal solvability)

Let F be a proper closed set-valued map from a Banach space X to a Banach space Y. We assume that there exists a constant $c > 0$ such that, for all $(x,y) \in \text{graph}\,(F)$, for all $v \in Y$, there exists $u \in X$ satisfying $v \in DF(x,y)(u)$ and $\|u\| \leq c\|y\|$. Then F maps X onto Y and its inverse F^{-1} is a Lipschitz set-valued map with Lipschitz constant equal to c.

▲

Let us mention also the following consequence of the proof of Theorem 4.

Corollary 7

Let F be a proper closed set-valued map from a Banach space X to a Banach space Y. Assume that there exists a constant $c > 0$ such that

$$\text{(18)} \qquad \begin{aligned} &\forall (x,y) \in \text{graph}\,(F),\ \exists u \in X \text{ satisfying} \\ &-y \in DF(x,y)(u) \text{ and } \|u\| \leq c\|y\| \quad . \end{aligned}$$

Then the set $F^{-1}(0)$ of zeros of F is nonempty and

$$\text{(19)} \qquad \forall x \in \text{Dom}\,(F), \quad d(x,F^{-1}(0)) \leq c\,d(0,F(x)) \quad .$$

Remark

When $F = f|_K$ is the restriction to a closed subset K of a continuously differentiable single-valued map f, assumption (18) becomes

$$\text{(20)} \qquad \begin{aligned} &\forall x \in K,\ \exists u \in T_K(x) \text{ such that } -f(x) = \nabla f(x)u \\ &\text{and } \|u\| \leq c\|f(x)\| \quad . \end{aligned}$$

We then deduce that there exists a solution $\bar{x} \in K$ to the equation $f(\bar{x}) = 0$ and that

(21) $\quad \forall x \in K, \quad d(x, f^{-1}(0)) \leq c\|x\| \quad .$

In the book by Aubin and Cellina (1983), it is shown that assumption (20) implies the existence of a trajectory of the implicit differential equation

(22) $$\begin{cases} \nabla f(x(t))x'(t) = -f(x(t)) \\ x(0) = x_0 \text{ given in } K \end{cases}$$

satisfying

(23) $\quad \forall t \geq 0, \quad x(t)$ belongs to K

and

(24) $\quad d(x(t), f^{-1}(0)) \leq e^{-ct}\, x(0) \quad .$

We observe that the differential equation (22) is the continuous version of the Newton method and that inequality (24) implies the convergence of the Newton method. ■

Remark

When the graph of F is compact (i.e., when the domain of F is compact and F is upper semicontinuous with compact values), we need only to assume that

(25) $\quad \forall (x,y) \in \text{graph } (F), \ \exists\, u \in X$ satisfying $-y \in DF(x,y)(u)$

for deducing that $F^{-1}(0)$ is nonempty.

Indeed, we minimize on the graph of F the function $(x,y) \to \|y\|$ and we denote by $(\bar{x},\bar{y}) \in \text{graph } (F)$ a minimizer. We proceed as in the proof of Theorem 4 with $\varepsilon = 0$.

■

The use of set-valued maps abolishes the formal distinction between the Inverse Function Theorem and the Implicit Function Theorem.

Let X, Y and Z be three Banach spaces and G be a set-valued map from $X \times Y$ to Z. The Implicit Function Theorem deals with the behavior of the map that associates to any $(y,z) \in Y \times Z$ the set of solutions x to the inclusion $z \in G(x,y)$. This amounts to study the inverse of the set-valued map F from X to $Y \times Z$ defined by

(26) $$(y,z) \in F(x) \iff z \in G(x,y) \quad .$$

Since the graphs of the set-valued maps F and G coincide as subsets of $X \times Y \times Z$, there are close relations between the derivatives of F and G at (x_0,y_0,z_0), since the graphs of these two derivatives coincide with the tangent cone to the graph of G at (x_0,y_0,z_0). Then we can state the Implicit Function Theorem.

Theorem 8

Let G be a proper set-valued map with closed graph from $X \times Y$ to Z and (x_0,y_0,z_0) belong to the graph of G. We assume that

i) Both X,Y and Z are finite dimensional

(27) ii) $\forall v,\ w \in Y \times Z,\ \exists u \in X$ such that

$$w \in CF(x_0,y_0,z_0)(u,v) \quad .$$

Then

(28) $$F^{-1} \text{ is pseudo-Lipschitz around } (x_0,(y_0,z_0)) \quad .$$

▲

In the case when G is a continuously differentiable function, we obtain the following useful corollary.

Corollary 9

Let g be a C^1 function from an open neighborhood of (x_0,y_0) in $X \times Y$ to Z satisfying

(29) $\nabla_y g(x_0,y_0)$ is surjective from Y to Z .

Then there exist neighborhoods V_0 and V_1 of y_0, $V_0 \subset V_1$, neighborhoods U of x_0 and W of z_0 and a constant $c > 0$ such that

(30) $\forall x \in U, \quad \forall z \in W, \exists y \in V_0$ such that $g(x,y) = z$

and, if we set $F^{-1}(x,z) := \{y \in Y | g(x,y) = z\}$,

(31)
$$\forall x_1, x_2 \in V \quad , \quad z_1, z_2 \in W \quad ,$$
$$d(F^{-1}(x_1,z_1) \cap V_0,\ F^{-1}(x_2,z_2) \cap V_1) \leq \ell(\|x_1-x_2\| + \|z_1-z_2\|)$$

▲

Proof

It is analogous to the proof of Corollary 5 and follows from Theorem 4 applied to the set-valued map F from Y to $X \times Z$ defined by

(32)
$$F(y) := \begin{cases} \{(x,z) \in K \times Z | g(x,y) = z\} \text{ when } y \in L \\ \emptyset \text{ when } y \notin L \end{cases}$$

where K and L are closed neighborhoods of z_0 and y_0 on which g is C^1. The graph of F is closed and assumption (7) is satisfied: let $(u,w) \in X \times Z$ be chosen and define $v \in Y$ as a solution to the equation

(33) $$\nabla_y g(x_0,y_0)v = w - \nabla_x g(x_0,y_0)u$$

satisfying

(34) $$\|v\| \leq c\|w - \nabla_x g(x_0,y_0)u\|$$

thanks to the Banach open mapping principle. We set $\hat{w} := \nabla g(x,y)(u,v) - \nabla g(x_0,y_0)(u,v)$. We see that

(35) $$(u,w) \in DF(y;(x,z))(v) + (0,\hat{w})$$

with

(36) $\|v\| \leq c \max (1, \|\nabla x g(x_0,y_0)\|) (\|u\| + \|w\|)$

and, α being given,

(37) $\|(0,\hat{w})\| \leq \|\nabla g(x,y) - \nabla g(x_0,y_0)\| (\|u\| + \|w\|) \leq \alpha(\|u\| + \|w\|)$

provided that (x,y) remains in a small neighborhood of (x_0,y_0). Hence Theorem 4 implies that F^{-1} is pseudo-Lipschitz around $(x_0,(y_0,z_0))$, which is what the conclusion of Corollary 9 states. ■

This being said, it is not always obvious to obtain "nice" formulas for the derivatives. For instance, let X and Y be Banach spaces, A be a continuous linear operator from X to Y, $G : X \to X^*$ and $H : Y \to Y^*$ be set-valued maps.

We consider the set of solutions $x \in X$ to the inclusion

(38) $p \in G(x) + A^*H(Ax+y)$,

where p is given in X^*. The first idea is to apply an Inverse Function Theorem to the set-valued map E from X to $X^* \times Y$ defined by

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

Unfortunately, there is no "nice" expression for the tangent cone to the graph of E in $X \times X^* \times Y$.

But we can introduce an auxiliary variable $q \in Y^*$ and write inclusion (38) as the equivalent inclusion

$$(40) \quad \begin{cases} \text{i)} & p \in G(x) + A^*q \\ \text{ii)} & y \in -Ax + H^{-1}(q) \end{cases}$$

The set of solutions (x,q) to this problem is denoted by $F^{-1}(p,y,A)$, where F is the set-valued map from $X \times X^*$ to $X^* \times Y \times L(X,Y)$ defined by

(41) $(p,y,A) \in F(x,q)$ if and only if (40) holds.

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

Lemma 10

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

$$(42)\quad \begin{cases} \text{i)} & p_0 \in G(x_0) + A_0^* q_0 \\ \text{ii)} & y_0 \in -Ax_0 + H^{-1}(q_0) \end{cases}$$

The following conditions are equivalent

$$(43)\quad (\delta p, \delta y, \delta A) \in CF(x_0, q_0; p_0, y_0, A_0)(\delta x, \delta q)$$

$$(44)\quad \begin{cases} \text{i)} & \delta p - \delta A^* \cdot q_0 \in CG(x_0, p_0 - A_0^* q_0)(\delta x) + A_0^* \delta q \\ \text{ii)} & \delta y + \delta A \cdot x_0 \in -A_0 \delta x + CH^{-1}(q_0, y_0 + Ax_0)(\delta q) \end{cases}$$

Proof

a) We prove that (43) implies (44). We choose sequences $(x_n, q_n, p_n', y_n', h_n)$ converging to $(x_0, q_0, p_0 - A^* q_0, y_0 + Ax_0, 0)$. 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 (43), 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

$$\text{i)}\quad p_n' + h_n(\delta p_n - A_0^* \delta q_n - \delta A_n^* q_n + h_n \delta A_n^* \delta q_n) \in G(x_n + h_n \delta x_n)$$

$$\text{ii)}\quad y_n' + h_n(\delta y_n + A_0 \delta x_n + \delta A_n \cdot x_n + h_n \delta A_n \delta x_n) \in H^{-1}(q_n + h_n \delta q_n) \quad .$$

Hence the system of inclusions (44) holds true.

b) Conversely, let us consider the system of inclusion (44) and let us prove (43). We choose sequences $(x_n, q_n, y_n, p_n, A_n, h_n)$ converging to $(x_0, q_0, y_0, A_0, 0)$. Then we know that there exist sequences of elements δx_n, δq_n, δu_n and v_n converging to δx, δq, $\delta p - \delta A^* q_0 - A_0^* \delta q$, and $\delta y + \delta A x_0 + A_0 \delta x$ respectively. We set

(45)

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) $\delta y_n := v_n - \delta A x_n - A_n \delta x_n - h_n \delta A x_n$, 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 (43) holds true.

Corollary 11

Let X and Y be finite dimensional and $G : X \to X^*$ and $H : Y \to Y^*$ be set-valued with closed graph. Let (p_0, y_0, A_0) belong to $X^* \times Y \times L(X,Y)$. Assume that there exists a solution (x_0, q_0) to the system of inclusions

(46)

i) $p_0 \in G(x_0) + A_0^* q_0$

ii) $y_0 \in -A_0 x_0 + H^{-1}(q_0)$.

If the matrix of closed convex processes

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

is surjective, then there exist neighborhoods U and V of (x_0, q_0),

$U \subset V$, and W of (p_0,y_0,A_0) such that the set-valued map

(47) $$(p,y,A) \in W \to F^{-1}(p,y,A) \cap U$$

has nonempty values and is pseudo-Lipschitz. Furthermore, the derivative of F^{-1} is given by

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

▲

6. CALCULUS OF CONTINGENT AND TANGENT CONES, DERIVATIVES AND EPI-DERIVATIVES

The applications of nonsmooth analysis to nonlinear analysis to which we have devoted the preceding section motivate the development of a calculus of contingent and tangent cones, of contingent derivatives and derivatives of set-valued maps and of epi-contingent derivatives and epi-derivatives of real-valued functions.

We summarize this calculus in the Appendix, adding the formulas of convex analysis for the sake of comparison.

<u>Proposition 1</u>

a) Let $K \subset L \subset X$ be two nonempty subsets. Then

(1) $$\forall x_0 \in K, \quad T_K(x_0) \subset T_L(x_0) \quad .$$

b) Let $K := \bigcup_{i\in I} K_i$ be the union of subsets K_i. Then

(2) $$\forall x_0 \in K, \quad \bigcup_{i\in I} T_{K_i}(x_0) \subset T_K(x_0) \quad .$$

If $I = \{1,\dots,n\}$ is finite, equality holds true.

c) Let $K := \bigcap_{i \in I} K_i$ be the intersection of subsets K_i. Then

(3) $$\forall x_0 \in K, \quad T_K(x_0) \subset \bigcap_{i \in I} T_{K_i}(x_0)$$

d) Let $K := \prod_{i \in I} K_i$ be a finite product of subsets K_i. Then

(4) $$\forall x_0 := \left(x_0^i\right)_{i \in I} \in K, \quad T_K(x_0) \subset \prod_{i=1}^{n} T_{K_i}(x_0^i) \text{ and } C_K(x_0) = \prod_{i \in I} C_{K_i}(x_0^i).$$

Proposition 2

Let X and Y be Banach spaces, A be a C^1-map from an open subset Ω of X to Y and $K \subset \Omega$ be a subset of X. Then

(5) $$\forall x_0 \in K, \quad \nabla A(x_0) T_K(x_0) \subset T_{A(K)}(A(x_0))$$

▲

Proof

Let v_0 belong to $T_K(x_0)$; there exist sequences $h_n \to 0+$ and $v_n \to v_0$ such that $x_0 + h_n v_n$ belongs to K for all n. Then the sequence of elements $u_n := (A(x_0+h_n v_n)-A(x_0))/h_n$ converges to $A(x_0)v_0$ and $A(x_0) + h_n u_n$ belongs to $A(K)$ for all n. Hence $A(x_0)v_0$ belongs to $T_{A(K)}(Ax_0)$. ■

In particular, if $A \in L(X,Y)$, we obtain the formula

(6) $$\forall x \in K, \quad AT_K(x) \subset T_{A(K)}(A(x)) \quad .$$

We now study the contingent cone to the preimage of a set by a smooth map:

Proposition 3

a) Let X and Y be two Banach spaces, $L \subset X$ and $M \subset Y$ be two subsets and A be a C^1-map from an open neighborhood of L to Y. We set

(7) $$K := \{x \in L \mid A(x) \in M\} = L \cap A^{-1}(M) \quad .$$

Then,

(8) $$\forall x \in K, \quad T_K(x) \subset T_L(x) \cap \nabla A(x)^{-1} T_M(A(x)) \quad .$$

b) Let X and Y be finite dimensional spaces, A be a C^1-map from an open subset $\Omega \subset X$ to Y, $L \subset \Omega$ and $M \subset Y$ be closed subsets of X and Y respectively.

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

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

Then

(10) $$C_L(x_0) \cap \nabla A(x_0)^{-1} C_M(Ax_0) \subset C_K(x_0) \quad .$$

Proof

a) By Proposition 1, $T_K(x) \subset T_L(x)$ because $K \subset L$. By Proposition 2, $\nabla A(x) T_K(x) \subset T_{A(K)}(Ax) \subset T_M(Ax)$ because $A(K) \subset M$. Hence $T_K(x) \subset \nabla A(x)^{-1} T_M(Ax)$ and consequently, formula (8) holds true.

b) We introduce the set-valued map F from X to Y defined by

(11) $$F(x) := A(x) - M \text{ when } x \in L, \ F(x) := \emptyset \text{ when } x \notin L \quad .$$

We observe that $F^{-1}(0) = K$. We shall prove that there exists a neighborhood U_0 of x_0 in L such that

(12) $$\forall x \in U_0, \quad d(x, F^{-1}(0)) \leq \ell d_M(Ax) \quad .$$

Indeed, 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 exist a neighborhood U of x_0, a ball of radius r in Y and a constant $\ell > 0$ such that:

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

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

$$\|A(x) - \pi_M(A(x))\| \le \|A(x) - A(x_0)\| < r$$

Therefore, we know that for all $x \in U_0 := L \cap U$,

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

c) 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 $U_0 := L \cap U$ when $\|x - x_0\| \le \alpha$ and $h \le \beta$. Since $F^{-1}(0) = L \cap A^{-1}(M)$, we deduce from (12)

$$\frac{d(x+hu_0, F^{-1}(0))}{h} \le c\, \frac{d_M(A(x+hu_0))}{h} \le \frac{d_M(Ax + h\nabla A(x_0)u_0)}{h}$$

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

The first term on the right-hand side converge 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 $C_{F^{-1}(0)}(u_0)$. ■

<u>Corollary 4</u>

Let X and Y be finite dimensional spaces, A be a continuously differentiable map from X to Y and M be a closed subset of Y. Let Ax_0 belong to M. If

(13) $\quad \mathrm{Im}\, \nabla A(x_0) - C_M(Ax_0) = Y$

then

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

Corollary 5

Let L and M be two nonempty closed subsets of a finite dimensional space X and x_0 belong to $L \cap M$. If

(15) $$C_L(x_0) - C_M(x_0) = X \quad ,$$

then

(16) $$C_L(x_0) \cap C_M(x_0) \subset C_{L\cap M}(x_0) \quad .$$

Corollary 6

Let K_i $(i=1,\dots,n)$ be n nonempty closed subsets of X and x_0 belong to $\bigcap_{i=1}^{n} K_i$. We posit the following assumption

(17) $$\forall v_1,\dots,v_n \in X, \quad \bigcap_{i=1}^{n} (C_{K_i}(x_0) - v_i) \neq \emptyset \quad .$$

Then

(18) $$\bigcap_{i=1}^{n} C_{K_i}(x_0) \subset C_{\bigcap_{i=1}^{n} K_i}(x_0) \quad .$$

Proof

Let $\vec{D} \subset X^n$ denote the closed vector space of constant sequences $\vec{x} := (x,\dots,x)$. Then K is identified with $\vec{K} := \vec{D} \cap \prod_{i=1}^{n} K_i$. We observe that $C_{\vec{D}}(\vec{x}) = \vec{D}$ and that $C_{\prod_{i=1}^{n} K_i}(\vec{x}) = \prod_{i=1}^{n} C_{K_i}(x)$. Assumption (17) implies that

(19) $$C_{\vec{D}}(\vec{x}) - C_{\prod_{i=1}^{n} K_i}(\vec{x}) = \vec{D} - \prod_{i=1}^{n} C_{K_i}(x) = X^n \quad .$$

Therefore Corollary 5 implies that

(20) $$\vec{D} \cap \prod_{i=1}^{n} C_{K_i}(x) \subset C_{\vec{D}} \cap \prod_{i=1}^{n} K_i(\vec{x}) \quad ,$$

i.e., inclusion (18). ■

We shall derive from the properties of the contingent and tangent cones a calculus of contingent derivatives and derivatives of set-valued maps.

<u>Proposition 7</u>

a) Let F be a set-valued map from X to Y and B be a C^1-map from an open neighborhood Ω of $\mathrm{Im}\, F \subset Y$ to Z.

Then

(21) $$\forall u_0 \in X, \quad \nabla B(y_0) \cdot DF(x_0,y_0)(u_0) \subset D(BF)(x_0,By_0)(u_0) \quad .$$

If

(22) F is Lipschitz around $x_0 \in \mathrm{Int\, Dom}\, F$ with compact values and $\dim Y < +\infty$ then

(23) $$\forall u_0 \in X, \quad \nabla B(y_0) \cdot DF(x_0,y_0)(u_0) = D(BF)(x_0,By_0)(u_0) \quad .$$

b) Let F be a set-valued map from X to Y and A be a C^1-map from X_0 to X. Then, if Ax_0 belongs to Dom F,

(24) $$\forall u_0 \in X_0, \quad D(FA)(x_0,y_0)(u_0) \subset DF(Ax_0,y_0)(\nabla A(x_0)(u_0)) \quad .$$

If we assume that either

(25)a) F is Lipschitz around Ax_0 ,

or

(25)b) $\nabla A(x_0)$ is surjective and $\dim X_0 < +\infty$,

then

(26) $$\forall u_0 \in X_0, \quad D(FA)(x_0,y_0)(u_0) = DF(Ax_0,y_0)(\nabla A(x_0)(u_0) \quad .$$

Proof

a) Let $(1 \times B)$ be the map: $(x,y) \in X \times \Omega \to (x,B(y)) \in Y \times Z$. The graph of the set-valued map $G := BF$ is related to the graph of F by the relation: graph $(G) = (1 \times B)$graph (F). By Proposition 2, we know that $(1 \times \nabla B(y_0))T_{graph(F)}(x_0,y_0)$ is contained in $T_{graph(G)}(x_0,By_0)$. This implies formula (21).

b) Let w_0 belong to $D(BF)(x_0,By_0)(u_0)$. There exist sequences $h_n \to 0$, $u_n \to u_0$ and $w_n \to w_0$ such that $h_n w_n$ belongs to $B(F(x_0+h_n u_n)) - B(y_0)$ for all n. Hence there exists $v_n \in (F(x_0+h_n u_n))-y_0)/h_n$ such that $h_n w_n = B(y_0+h_n v_n) - B(y_0)$. Since F is Lipschitz around x_0, v_n belongs to $F(x_0) - y_0 + h_n \| u_n \| B$, which is contained in a compact set, because the values of F are compact and the dimension of Y is finite. Hence a subsequence (again denoted by) v_n converges to some v_0, which belongs to $DF(x_0,y_0)(u_0)$, and thus, the sequence of elements w_n converges to $\nabla B(y_0) \cdot v_0 = w_0$. Therefore $D(BF)(x_0,By_0)(u_0) \subset \nabla B(y_0) DF(x_0,y_0)(u_0)$.

c) Let $A \times 1 : X_0 \times Y \to X \times Y$ be the map defined by $(A \times 1)\ (x_0,y) = (Ax_0,y)$. The graph of the map $G = FA$ is related to the graph of F by the formula graph $G = (A \times 1)^{-1}$graph F. By Corollary 4, we know that $T_{graph(G)}(x_0,y_0) \subset (\nabla A(x_0) \times 1)^{-1} T_{graph(F)}(Ax_0,y_0)$. This implies formula (24).

d) To prove (26) let us pick w_0 in $DF(Ax_0,y_0)(\nabla A(x_0)u_0)$. There exist sequences $h_n \to 0+$, $v_n \to \nabla A(x_0)u_0$ and $w_n \to w_0$ such that $h_n w_n$ belongs to $F(A(x_0)+h_n v_n) - y_0$. Assume first that F is Lipschitz around x_0, we deduce that

$$F(A(x_0)+h_n v_n)) \subset F(A(x_0+h_n u_n)) + \ell \| A(x_0+h_n u_0)-A(x_0)-h_n v_n \| B \quad .$$

Hence there exist a sequence of elements w'_n converging to w_0 such that $h_n w'_n \in F(A(x_0+h_n u_0))-y_0)$ for all n. Thus w_0 belongs to $D(FA)(x_0,y_0)(u_0)$. Assume now that $\nabla A(x_0)$ is surjective. Then Corollary 5.5 implies the existence of a constant $\ell > 0$ such that, for n large enough, there exists a solution x_n to the equation $A(x_n) = A(x_0) + h_n v_n$ satisfying $\|x_n - x_0\| \le \ell h_n \, v_n$. Since $\dim x_0 < +\infty$, we deduce that a subsequence of elements $u_n := \frac{x_n - x_0}{h_n}$ converges to some element u_0; since $y_0 + h_n w_n \in FA(x_0+h_n u_n)$ for all $n \ge 0$, we deduce that w_0 belongs to $D(FA)(x_0,y_0)(u_0)$. ■

We now investigate the chain rule formula for the derivative.

Proposition 8

Let X_0, X and Y be three finite-dimensional spaces, F be a set-valued map with closed graph from X to Y and A be a continuously differentiable map from X_0 to X. Let x_0 belong to $A^{-1}\mathrm{Dom}\, F$ and $y_0 \in F(x_0)$. We assume that

(27) $$\mathrm{Im}\, \nabla A(x_0) - \mathrm{Dom}\, CF(Ax_0,y_0) = X \quad .$$

Then

(28) $$\begin{cases} \text{i)} \ \forall u_0 \in X_0, \ C(FA)(x_0,y_0)(u_0) \supset CF(Ax_0,y_0)(\nabla A(x_0)u_0) \\ \text{ii)} \ \forall q_0 \in Y^*, \ C(FA)(x_0,y_0)^*(q_0) \subset \nabla A(x_0)^* CF(Ax_0,y_0)^*(q_0) \end{cases}$$

Proof

Since graph (FA) = $(A \times 1)^{-1}$graph (F), we apply the second part of Proposition 3, which states that $(\nabla A(x_0) \times 1)^{-1} C_{\mathrm{graph}(F)}(Ax_0,y_0)$ is contained in $C_{\mathrm{graph}(FA)}(x_0,y_0)$, i.e., formula (28), provided that property

(29) $$\mathrm{Im}\, (\nabla A(x_0) \times 1) - C_{\mathrm{graph}(F)}(Ax_0,y_0) = X \times Y$$

is satisfied. But it follows from Assumption (27), which also implies

$$(CF(Ax_0,y_0) \cdot \nabla A(x_0))^* = \nabla A(x_0)^* CF(Ax_0,y_0)^* \quad .$$

Proposition 9

Let F be a proper set-valued map from X to Y, K be a subset of X and x_0 belong to $K \cap \mathrm{Dom}\, F$. Then

(30) $$D(F|_K)(x_0,y_0)(u_0) \subset DF(x_0,y_0)|_{T_K(x_0)}(u_0) \quad .$$

If X and Y are finite dimensional, the graph of F is closed, K is closed and

(31) $$C_K(x_0) - \mathrm{Dom}\, CF(x_0,y_0) = X \quad ,$$

then

(32) $$CF(x_0,y_0)|_{C_K(x_0)}(u_0) \subset C(F|_K)(x_0,y_0)(u_0)$$

and, for all $q_0 \in Y^*$,

(33) $$C(F|_K)(x_0,y_0)^*(q_0) \subset CF(x_0,y_0)^*(q_0) + N_K(x_0) \quad .$$

▲

Proof

We observe that graph $(F|_K)$ = graph $(F) \cap (K \times Y)$. Then $T_{\mathrm{graph}(F|_K)}(x_0,y_0) \subset T_{\mathrm{graph}(F)}(x_0,y_0) \cap (T_K(x_0) \times Y)$, from which we deduce formula (30). We observe that Assumption (31) implies that

$$C_{\mathrm{graph}(F)}(x_0,y_0) - C_K(x_0) \times Y = X \times Y \quad .$$

Therefore, Corollary 5 implies that

$$C_{\text{graph}(F)}(x_0,y_0) \cap (C_K(x_0) \times Y) \subset C_{\text{graph}(F|_K)}(x_0,y_0)$$

from which we deduce formula (32). Assumption (31) allows us to deduce formula (33) from formula (32).

■

Proposition 10

a) Let us consider n set-valued maps F_i from X to Y.

$$(34) \qquad \forall u_0 \in X, \quad D \bigcup_{i=1}^{n} F_i \; (x_0,y_0)(u_0) = \bigcup_{i=1}^{n} DF_i(x_0,y_0)(u_0)$$

b) Let us consider n set-valued maps F_i with closed graph from a finite dimensional space X to a finite dimensional space Y. Let (x_0,y_0) belong to the intersection of the graphs of F_i. Assume that

$$(35) \qquad \begin{cases} \forall (u_i,v_i) \in X \times Y \quad (i=1,\dots,n), \quad \exists (u_0,v_0) \in X \times Y \\ \text{such that} \\ v_0 \in CF_i(x_0,y_0)(u_0+u_i) - v_i \quad \text{for} \quad i=1,\dots,n \end{cases} .$$

Then

$$(36) \qquad \forall u_0 \in X, \quad C\left(\bigcap_{i=1}^{n} F_i\right)(x_0,y_0) \supset \bigcap_{i=1}^{n} CF_i(x_0,y_0)(u_0) \quad .$$

▲

Proof

We note that graph$(\cup F_i)$ = $\cup$graph (F_i) and graph $(\cap F_i)$ = $\cap$graph (F_i) and we apply Proposition 3 and Corollary 6 respectively.

■

We deduce at once a calculus of epi-contingent derivatives and epi-derivatives of real-valued functions V from the calculus of contingent derivatives and derivatives of the set-valued map

$\underset{\sim}{V}_+$ defined by $\underset{\sim}{V}_+(x) := V(x) + \mathbb{R}_+$ when $V(x) < +\infty$ and $\underset{\sim}{V}_+(x) := \emptyset$ when $V(x) = \{+\infty\}$.

Proposition 11

a) Let V be a proper function from Y to $\mathbb{R} \cup \{+\infty\}$ and A a C^1-map from X to Y. If Ax_0 belongs to the domain of V, then

(37) $$\forall u_0 \in X, \quad D_+V(Ax_0)(\nabla A(x_0)(u_0) \leq D_+(V \cdot A)(x_0)(u_0)$$

b) Let X and Y be finite-dimensional spaces, A be a C^1-map from X to Y and $V : Y \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function. Let x_0 belong to $A^{-1}(\mathrm{Dom}\, V)$. If we assume that

(38) $$\mathrm{Im}\, \nabla A(x_0) - \mathrm{Dom}\, C_+V(Ax_0) = Y \quad ,$$

then

(39) $$\forall u_0 \in X, \quad C_+V(Ax_0)(\nabla A(x_0)u_0) \geq C_+(V \cdot A)(x_0)(u_0)$$

and

(40) $$\partial(V \cdot A)(x_0) \subset \nabla A(x_0)^* \partial V(Ax_0) \quad .$$

Observe that Assumption (38) is satisfied when either V is Lipschitz around $A(x_0)$, because in this case $\mathrm{Dom}\, C_+V(Ax_0) = Y$, or $\nabla A(x_0)$ is surjective.

Proposition 9 implies the following formula for epi-derivatives of restrictions.

Proposition 12

Let X be a finite-dimensional space, V a proper lower semicontinuous function from X to $\mathbb{R} \cup \{+\infty\}$ and $K \subset X$ a closed subset. Let x_0 belong to $K \cap \mathrm{Dom}\, V$. We assume that

(41) $$\mathrm{Dom}\, C_+V(x_0) - C_K(x_0) = X \quad .$$

Then

(42) $\forall u \in C_K(x_0), \quad C_+(V|_K)(x_0)(u) \leq C_+V(x_0)(u)$

and

(43) $\partial(V|_K)(x_0) \subset \partial V(x_0) + N_K(x_0) \quad .$

In particular, these formulas hold true when V is Lipschitz around x_0. ▲

Proposition 13

Let V be a proper function from the product $X \times Y$ of two Banach spaces to $\mathbb{R} \cup \{+\infty\}$. We set

(44) $$W(y) := \inf_{x \in X} V(x,y) \quad .$$

If $x_y \in X$ minimizes $x \to V(x,y)$ on X, then

(45) $$\forall v \in Y, \quad D_+W(y)(v) \leq \inf_{u \in X} D_+V(x_y,y)(u,v)$$

▲

Proof

Let u and v belong to X and Y respectively. Then the following inequalities hold true:

$$\frac{W(y+hv) - W(y)}{h} \leq \frac{V(x_y+hu,y+hv) - V(x_y,y)}{h} \quad .$$

This implies that for all (u_0,v_0) in $X \times Y$,

$$D_+W(y)(v_0) \leq D_+V(x_y,y)(u_0,v_0)$$ ■

and consequently, inequality (45).

Remark

When v = 0, we obtain Proposition 3.5 as a consequence. The analogous statement holds true for the supremum of a family of functions. Let

(46) $$U(y) := \sup_{x \in X} V(x,y) = V(\tilde{x}_y, y) \quad .$$

Then

(47) $$\forall v \in Y, \quad D_+U(y)(v) \geq \sup_{u \in X} D_+V(\tilde{x}_y, y)(u,v) \quad .$$ ∎

We now provide a formula on the epi-derivative of a supremum of a finite number of functions.

Proposition 14

Let us consider n proper lower semicontinuous functions from a finite dimensional space X to $\mathbb{R} \cup \{+\infty\}$. We set $U(x) := \max_{i=1,\ldots,n} V_i(x)$. We assume that

(48) $$x_0 \text{ belongs to } \bigcap_{i=1}^{n} \operatorname{Int} \operatorname{Dom} V_i \quad .$$

Let us set $J(x_0) := \{i=1,\ldots,n \mid V_i(x_0) = U(x_0)\}$. Assume also that

(49) $$\begin{cases} \forall i \in J(x_0), \quad \forall u_i \in X, \text{ there exists } u_0 \\ \text{such that } u_0 - u_i \in \operatorname{Dom} C_+V_i(x_0) \text{ for all } i \in J(x_0) \end{cases}$$

Then

(50) $$\forall u_0 \in X, \quad C_+U(x_0)(u_0) \leq \max_{i \in J(x_0)} C_+V_i(x_0)(u_0)$$

and

(51) $$\partial U(x_0) \subset \overline{\text{co}} \bigcup_{i \in J(x_0)} \partial V_i(x_0) \quad .$$

Proof

Assumption (48) implies that when $i \notin J(x_0)$, then $(x_0, U(x_0))$ belongs to the interior of $\text{Ep}V_i$, so that the tangent cone $C_{\text{Ep}(V_i)}(x_0, U(x_0))$ is equal to $X \times \mathbb{R}$. Assumption (49) implies that

(52) $\forall (u_i, \lambda_i) \in X \times \mathbb{R}$ (i=1,...,n), there exists (u_0, λ_0) such that $(u_0 - u_i, \lambda_0 - \lambda_i)$ belongs to $C_{\mathcal{E}p(V_i)}(x_0, U(x_0))$ for all i=1,...,n .

Indeed, u_0 is given by property (49) and we take

$$\lambda_0 := \max_{i \in J(x_0)} (C_+ V_i(x_0)(u_0 - u_i) + \lambda_i) \quad .$$

Therefore, we are allowed to use Corollary 6, since $\mathcal{E}pU = \bigcap_{i=0}^{n} \mathcal{E}p(V_i)$. Then

$$\bigcap_{i \in J(x_0)} C_{\mathcal{E}p(V_i)}(x_0, U(x_0)) = \bigcap_{i=1}^{n} C_{\mathcal{E}p(V_i)}(x_0, U(x_0)) \subset C_{\mathcal{E}pV}(x_0, U(x_0)) \quad .$$

This inclusion implies inequality (50), from which we deduce inclusion (51), because the support function of a union is the supremum of the support functions. ■

We mention the following corollary.

<u>Corollary 15</u>

Assume that the n proper lower semicontinuous functions are Lipschitz around x_0. Then

(53) $$\forall u_0 \in X, \quad C_+ U(x_0)(u_0) \leq \max_{i \in J(x_0)} C_+ V_i(x_0)(u_0)$$

and

(54) $$\partial U(x_0) \subset \overline{co} \bigcup_{i \in J(x_0)} \partial V_i(x_0) \quad .$$

▲

We now turn our attention to the behavior of epi-contingent derivatives and epi-derivatives of the sum of two functions. This time, we need the concept of strict epi-differentiability.

Proposition 16

Let V and W be proper functions from X to $\mathbb{R} \cup \{+\infty\}$. Then

$$(55) \qquad D_+V(x_0)(u_0) + D_+W(x_0)(u_0) \le D_+(V+W)(x_0)(u_0) \quad .$$

Assume that W is strictly epi-differentiable at x_0 and that

$$(56) \qquad \mathrm{Dom}\, C_+V(x_0) \cap \mathrm{Dom}\, B_+W(x_0) \neq \emptyset \quad .$$

Then

$$(57) \qquad C_+(V+W)(x_0)(u_0) \le C_+V(x_0)(u_0) + C_+W(x_0)(u_0)$$

and

$$(58) \qquad \partial(V+W)(x_0) \subset \partial V(x_0) + \partial W(x_0) \quad .$$

▲

Remark

Assumption (56) is satisfied when W is Lipschitz around x_0.

Proof

The formula for epi-contingent derivatives is obvious. We observe that for all u_0 belonging to $\mathrm{Dom}\, C_+V(x_0) \cap \mathrm{Dom}\, B_+W(x_0)$, we have

$$(59) \qquad C_+(V+W)(x_0)(u_0) \le C_+V(x_0)(u_0) + B_+W(x_0)(u_0) \quad .$$

Now, let $u \in \mathrm{Dom}\, C_+V(x_0) \cap \mathrm{Dom}\, C_+W(x_0)$ and $\lambda \in]0,1[$ be fixed. Since $\mathrm{Dom}\, B_+W(x_0) = \mathrm{Int}\, \mathrm{Dom}\, C_+W(x_0)$, we deduce that $(1-\lambda)u + \lambda u_0$ belongs to $\mathrm{Dom}\, C_+V(x_0) \cap \mathrm{Dom}\, B_+W(x_0)$, so that formula (59) implies that

$$\begin{aligned} C_+(V+W)(x_0)((1-\lambda)u+\lambda u_0) &\le (1-\lambda)C_+V(x_0)(u) + \lambda C_+V(x_0)(u_0) \\ &\quad + B_+W(x_0)((1-\lambda)u+\lambda u_0) \le (1-\lambda)C_+V(x_0)(u) + \lambda C_+V(x_0)(u_0) \\ &\quad + (1-\lambda)C_+W(x_0)(u) + \lambda B_+W(x_0)(u_0) \end{aligned}$$

by formula (27) of Section 3.

By letting λ converge to 0, we deduce formula (57), which implies formula (58), because the support function of a sum is the sum of support functions.

<u>Remark</u>

Observe that Assumption (56) implies that

$$\text{(60)} \qquad \text{Dom } C_+V(x_0) - \text{Dom } C_+W(x_0) = X \quad .$$

In the same way, we can prove without using the Inverse Function Theorem that the assumption

$$\text{(61)} \qquad \text{Dom } C_+V(x_0) \cap \nabla A(x_0)^{-1}\text{Dom } B_+W(x_0) \neq \emptyset$$

implies that

$$\text{(62)} \qquad C_+(V + W \quad A)(x_0)(u_0) \leq C_+V(x_0)(u_0) + C_+W(Ax_0)(Ax_0)(\nabla A(x_0)u_0)$$

and that

$$\text{(63)} \qquad \partial(V + W \quad A)(x_0) \subset \partial V(x_0) + \nabla A(x_0)^*\partial W(Ax_0)$$

without assuming that the spaces are finite dimensional.

■

7. COMMENTS

Nonsmooth analysis started at the end of the sixties for extending the successful subdifferential calculus to nonconvex and nonsmooth functions or for using convenient "tangent cones" for expressing the necessary conditions. Let us quote, among many other works, Dubovickii and Miljutin (1971), Ioffe and Tihomirov (1972), Neustadt (1976), etc.

The concept of generalized gradient and normal cone introduced by Clarke (1975) gave a new impetus in the field and was at the origin of a considerable amount of work. Other attempts for defining other concepts of generalized gradients were made by Russian mathematicians, see for instance Demianov-Vassiliev

(1981) and Pchenichny (1980). The part of this chapter dealing with generalized gradient and normal cones is based on the works of Clarke (1975), (1976d), (1977b), (1981a), regrouped in Clarke (1983), and the works of Rockafellar (1979a,b,c) and (1980). The importance of the role of the Bouligand tangent cone (Bouligand, 1932) in viability theory for differential inclusion is stressed in the book by Aubin and Cellina (1984). The fact that the tangent cone is the Kuratowski liminf of the contingent cone was discovered by Cornet (1981), Penot (1981) and Rockafellar and Wets (unpublished). Many works were devoted to tangent cones and derivatives. (Auslender, 1978a,b; Crouzeix, 1977 (for quasi-convex functions), Frankowska, T.A. d); Gauvin, 1979; Gollan, 1981; Halkin, 1976, Hiriart-Urruty, 1978, 1979a,b,c; Hiriart-Urruty and Thibault, 1980; Hogan, 1973; Ioffe, 1981a, T.A. a,b; Janin, 1982; Lebourg, 1975, 1979; Lemaréchal, 1975; Penot, 1974, 1978a,b,c; Shi Shu Chung, 1980; Thibault, 1979; Warga, 1976, 1978a; Watkins,

Many concepts of generalized derivatives of vector-valued maps have been proposed and studied. Let us mention the fans, introduced by Ioffe, 1979, 1982 (see also Aubin, 1982b), 1981c, Hiriart-Urruty T.A.b, Kutateladze, 1977, Sweester, 1977, Thibault, 1982, T.A.a.

The concept of contingent derivative of set-valued map was introduced in Aubin (1981) and the concept of derivative in Aubin (1982a).

Other concepts of derivatives of set-valued maps were proposed by de Blasi (1976), Boudourides and Shinas (1981), Merica (1980), Nurminski (1978), Penot (T.A.a), Petcherskaja (1980), Shinas and Boudourides (1981), Thibault (T.A.b).

The inverse function theorem is taken from Aubin (1982a). See also the paper of Clarke (1976), Halkin (1976), Ioffe (1981c), Warga (1978) and the recent works of Rockafellar (unpublished). Epi-contingent derivatives are quite useful for the theory of Hamilton-Jacobi equations (see Aubin, (1981) and Aubin and Cellina, (1984)) and are related to the concept of generalized solutions introduced in Lions P.L. (1982).

REFERENCES

Aubin, J.-P.

1978 Gradients généralisés de Clarke. Ann. Sc. Math. Québec 2, 197-252.

1980 Further properties of Lagrange multipliers in nonsmooth optimization. Appl. Math. Optim. 6, 79-90.

1981 Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions. Advances in Mathematics, Suppl. Studies, L. Nachbin (Ed.), Academic Press, 160-232.

1982a) Comportement Lipschitzien des solutions de problèmes de minimization convexes. C.R.A.S. 295, 235-238.

b) Ioffe's fans and generalized derivatives of vector-valued maps. In "Convex Analysis and Optimization", Aubin and Vinter (Eds.), Pitman, Boston.

1983 Lipschitz behavior of solutions to convex minimization problems. Math. Oper. Res.

Aubin, J.-P. and A. Cellina

1984 Differential Inclusions. Springer-Verlag

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

1977 Multiplicateurs de Lagrange en optimisation non convexe et applications. J. London Math. Soc. 16, 357-366.

1979 Shadow prices and duality for a class of optimal control problems. SIAM J. Opt. Contr. 17, 567-586.

Aubin, J.-P. and I. Ekeland

1984 Applied Nonlinear Analysis. Wiley, Interscience.

Aubin, J.-P. and R. Vinter, Ed.

1982 Convex Analysis and Optimization. Pitman, Boston.

Auslender, A.

1976 Optimisation: méthodes numériques. Masson, Paris

Auslender, A.

1977 Minimisation sans contraintes de fonctions localement lipschitziennes. C.R.A.S. 284, 959-961.

1978a) Differentiable stability in nonconvex and nondifferentiable programming, in Math. Progr. Study No. 10, P. Huard, Ed.

1978b) Stabilité différentiable en programmation non convexe non différentiable, C.R. Acad. Sci. Paris, 286, 575-577.

Borwein, J.

1978 Weak tangent cones and optimization in Banach spaces. SIAM J. Contr. Opt. 16, 512-522.

Bouligand, G.

1932 Introduction à la géometrie infinitésimale directe. Gauthier-Villars.

Boudourides, M. and J. Shinas

1981 The mean-value theorem for multifunctions. Bull. Math. Roum. 25, 128-141.

Chaney, R.W.

1982 Second-order sufficiency conditions for non differentiable programming problems. SIAM J. Contr. Opt. 20, 20-33.

Clarke, F.H.

1975a) Generalized gradients and applications. Trans. Am. Math. Soc. 205, 247-262.

1975b) The Euler-Lagrange differential inclusion. J. Diff. Equ. 19, 80-90.

1976a) Optimal solutions to differential inclusions. J. Opt. Th. Appl. 19, 469-478.

1976b) The maximum principle under minimal hypotheses. SIAM J. Contr. Opt. 14, 1078-1091.

Clarke, F.H.

1976c) The generalized problem of Bolza. SIAM J. Contr. Opt. 14, 682-699.

1976d) A new approach to Lagrange multipliers. Math. Oper. Res. 1, 165-174.

1976e) On the inverse function theorem. Pacific J. Math. 64, 97-102.

1977a) Extremal arcs and extended Hamiltonian systems. Trans. Am. Math. Soc. 231, 349-367.

1977b) Multiple integrals of Lipschitz functions in the calculus of variations. Proc. Am. Math. Soc. 64, 260-264.

1978a) Pointwise contraction criteria for the existence of fixed points. Bull. Canadian Math. Soc. 21, 7-11.

1978b) Nonsmooth analysis and optimization, in the Proceedings of the Internat. Congress of Math., Helsinki.

1978c) Solutions periodiques des equations Hamiltoniennes. C.R.A.S. 287, 951-957.

1979 Optimal control and the true Hamiltonian. SIAM Review 21, 157-166.

1980 The Erdmann condition and Hamiltonian inclusions in optimal control and the calculus of variations. Canadian J. Math. 32, 494-509.

1981a) Generalized gradients of Lipschitz functionals. Advances in Mathematics 40, 52-67. (MRC Tech. Rep. 1687, August 1976).

1981b) Periodic solutions of Hamiltonian inclusions. J. Diff. Equ. 40, 1-6.

1982 On Hamiltonian flows and symptotic transformations. SIAM J. Contr. Opt. 20, 355-359.

1983 Optimization and Nonsmooth Analysis. Wiley, Interscience, New York.

Cornet, B.

1981 Regular properties of tangent and normal cones. Cahiers de Mathématiques de la Décision. No. 8130.

Crouzeix, J.P.

1977 Contribution à l'étude des fonctions quasi-convexes. Thèse. Université de Clermont.

1980 On second-order conditions for quasi-convexity. Math. Progr. 18, 349-352.

De Blasi, F.S.

1976 On the differentiability of multifunctions. Pacific J. Math. 66, 67-81.

De Blasi, F.S. and J. Myjak

T.A. Some generic properties in convex and nonconvex optimization theory.

Demianov, V.F. and L.V. Vassiliev

1981 Nondifferentiable optimization (in Russian). Nauka, Moscow.

Demianov, V.F. and A.M. Rubinov

1980 On quasidifferentiable functionals. Dokl. A.N.SSR 250 (Soviet Math. Dokl., 21, 14-17).

Dubovickii, A.I. and Miljutin, A.A.

1971 Necessary conditions for a weak extremum in the general problem of optimal control (in Russian). Nauka, Moscow.

Fenchel, W.

1949 On conjugate convex functions. Canadian J. Math. 1, 73-77.

1951 Convex Cones, Sets and Functions, mimeographed lecture notes, Princeton University, Princeton, New Jersey.

1952 A remark on convex sets and polarity, Medd. Lunds Univ. Mat. Sem. (Supplementband, 82-89).

1956 Über konvexe Funktionen mit vorgeschriebenen Niveau-mannigfaltigkeiten. Math. Z. 63, 496-506.

1967 Proceedings of the Colloquium on Convexity. Copenhagen, 1965, Matematisk Institut, Copenhagen.

Frankowska, H.

T.A. a) First order necessary conditions in nonsmooth variational and control problems. SIAM J. Contr. Opt.

b) Necessary conditions for Bolza problems.

c) Hamiltonian inclusions for differential inclusion problems.

d) Adjoint differential inclusions associated with an optimal trajectory of a differential inclusion.

Gautier, S.

T.A. Differentiabilité des multiapplications

Gauvin, J.

1973 Fermés invariants par un système dynamique. C.R.A.S. 276, 1457-1460.

1979 The generalized gradient of a marginal function in mathematical programming. Math. Op. Res. 4, 458-463.

Gollan, B.

1981 Higher order necessary conditions for an abstract optimization problem, in Mathematical Programming Study 14, 69-76.

T.A. a) On the marginal function in nonlinear programming.

b) A general perturbation theory for abstract optimization problems. J. Opt. Th. Appl.

Halkin, H.

1972 Extremal properties of biconvex contingent equations, in Ordinary Differential Equations (NRL-MRC Conference), Academic Press.

1976 Mathematical programming without differentiability, in Calculus of Variations and Control Theory, D.L. Russell (ed.), Academic Press, 279-288.

1976 Interior mapping theorem with set-valued derivative. J. d'Analyse Math. 30, 200-207.

Hiriart-Urruty, J.B.

1978 Gradients généralisés de fonctions marginales. SIAM J. Contr. Opt. 16, 301-316.

1979a) Tangent cones, generalized gradients and mathematical programming in Banach spaces. Math. Oper. Res. 4, 79-97.

1979b) Refinements of necessary optimality conditions in nondifferentiable programming I. Appl. Math. and Opt., 63-82.

1979c) New concepts in nondifferentiable programming. Bull. Soc. Math. France, Memoire 60, 57-85.

1980 Mean value theorems in nonsmooth analysis. Numer. Funct. Anal. and Opt. 2, 1-30.

1981 Optimality conditions for discrete nonlinear norm-approximation problems, in Optimization and Optimal Control (Proc. Conf., Math. Res. Inst., Oberwolfach, 1980), pp. 29-41, Lecture Notes in Control and Information Sci. 30, Springer-Verlag, Berlin.

T.A.a) Refinements of necessary optimality conditions in nondifferentiable programming II. Math. Progr. Study, to appear.

b) Characterizations of the plenary hull of the generalized Jacobian matrix, to appear.

c) A better insight into the generalized gradient of the absolute value of a function, to appear.

Hiriart-Urruty, J.B. and L. Thibault

1980 Existence et charactérisation de différentielles généralisées. C.R.A.S. 290, 1091-1094.

Hogan, W.

1973 Directional derivatives for extremal value functions with applications to the completely convex case. Oper. Res. 21, 188-209.

Ioffe, A.D.

1976 An existence theorem for a general Bolza problem. SIAM J. Contr. Opt. 14, 458-466.

Ioffe, A.D.

1977 On lower semicontinuity of integral functionals I. SIAM J. Contr. Opt. 15, 521-538.

1978 Survey of measurable selection theorems: Russian literature supplement. SIAM J. Contr. Opt. 16, 728-733.

1979 Différentielles généralisées d'applications localement lipschitziennes d'un espace de Banach dans un autre. C.R.A.S. 289, 637-639.

1981a) Sous-différentielles approchées de fonctions numériques. C.R.A.S. 292, 675-678.

1981b) A new proof of the equivalence of the Hahn-Banach extension and the least upper bound properties. Proc. Amer. Math. Soc. 82, 385-390.

1981c) Nonsmooth analysis: differential calculus of non-differentiable mappings. Transamer. Math. Soc. 266, 1-56.

1982 Nonsmooth analysis and the theory of fans. In Convex Analysis and Optimization, Aubin J.P. and Vinter, R. (Eds.), Pitman, Boston, 93-117.

T.A.a) Calculus of Dine derivatives of functions and contingent coderivatives of set-valued maps.

b) Approximate subdifferential of nonconvex functions. Transamer. Math. Soc.

Ioffe, A.D. and V.L. Levin

1972 Subdifferentials of convex functions. Trans. Moscow Math. Soc. 26, 1-72.

Ioffe, A.D. and V.M. Tihomirov

1974 The theory of extremal problems. Nauka, Moscow (English translation, North-Holland, 1979).

Janin, R.

1982 Sur des multiapplications qui ont des gradients généralisés. C.R.A.S. 294, 115-117.

Kutaladze, S.S.

1977 Subdifferentials of convex operators. J. Math. Sibirsk 13, 1057-1064. In Russian.

Lebourg, G.

1975 Valeur moyenne pour un gradient généralisé. C.R.A.S. 281, 795-797.

1979 Generic differentiability of Lipschitzian functions. Trans. Amer. Math. Soc. 256, 125-144.

Lemaréchal, C.

1975 An extension of Davidon methods to nondifferentiable problems. Math. Programming Study. Nondifferentiability Optimization, 95-109. North-Holland.

Lions, P.L.

1982 Generalized solutions of Hamilton-Jacobi equations. Pitman, Boston.

Martelli, M. and A. Vignoli

1974 On differentiability of multivalued maps. Boll. Un. Mat. Ital. 10, 701-712.

Maurer, H.

1979a) Differential stability in optimal control problems. Appl. Math. Opt. 5, 283-295.

1979b) First order sensitivity of the optimal value function in mathematical programming and optimal control. In Proceedings of the Symposium on Mathematical Programming with Data Perturbations, Washington, D.C.

Mirica, S.

1980 A note on the generalized differentiability of mappings. Nonlinear Analysis, TAM, 4, 567-575.

T.A. The contingent cone and the in generalizing derivatives for vector-valued and set valued

Neustadt, L.

1976 Optimization. Princeton University Press.

Nurminski, E.A.

1978 On differentiability of multifunctions, Kibernetica 6, 46-48 (in Russian).

Panagiotopoulos, P.D.

1980 Superpotentials in the sense of Clarke and in the sense of Warga and applications, ZAMM 62.

T.A. Optimal control and parameter identification of structures with convex or non-convex strain energy density. Solid Mech. Archives, to appear.

Papageorgiou, N.S.

T.A. Nonsmooth analysis on partially ordered vector spaces. Part 2: nonconvex case, Clarke's theory, Pacific J. Math.

Pchenitchny, B.W.

1971 Necessary conditions for an extremum . Dekker Inc.

1980 Convex analysis and extremal problems (in Russian). Nauka-Moscow.

Penot, J.P.

1973 Calcul différentiel dans les espaces vectoriels topologiques. Studia Math. 47, 1-23.

1974 Sous-différentiels de fonctions numériques non convexes. C.R.A.S., Paris, 278, 1553-1555.

1978a) Calcul sous-différentiel et optimisation. J. Func. Analysis 27, 248-276.

1978b) Utilisation des sous-différentiels généralisés en optimisation. First meeting A.F.C.E.T. - S.M.F. on applied mathematics, Palaiseau, A.F.C.E.T. - S.M.F. Paris.

Penot, J.P.

1978c) The use of generalized subdifferential calculus in optimization theory. Proc. Third Symp. on Oper. Res. Mannheim, Methods of Operations Research 31, 495-511, Athenäum, Berlin.

1979 Fixed point theory without convexity. Bull. Soc. Math. Fr. Mem. 60.

1981 A characterization of tangential regularity. Nonlinear Analysis. T.A.M. 5, 625-643.

T.A.a) Differentiability of relations and differential stability of perturbed optimization problems.

b) On regularity conditions in mathematical programming.

Petcherskaja, N.A.

1980 On differentiability of set-valued maps. Vestnik Leningrad University (in Russian), 115-117.

Rockafellar, R.T.

1966 Characterization of the subdifferentials of convex functions. Pacific J. Math. 17, 497-510.

1967a) Conjugates and Legendre transforms of convex functions. Canad. J. Math. 19, 200-205.

1967b) Monotone processes of convex and concave type.

1968a) Integrals which are convex functionals. Pacific J. Math. 24, 525-540.

1968b) Duality in Nonlinear Programming. Mathematics of the Decision Sciences, Part 1, G.B. Dantzig and A.F. Veinott (eds.), Lectures in Applied Math., Vol. II, Amer. Math. Soc., 401-422.

1969 Measurable dependence of convex sets and functions on parameters. J. Math. Anal. Appl. 28, 4-25.

1970a) Convex Analysis, Vol. 28 of Princeton Math. Series, Princeton Univ. Press.

1970b) Conjugate convex functions in optimal control and the calculus of variations. J. Math. Anal. Appl. 332, 174-222.

1970c) Generalized Hamiltonian equations for convex problems of Lagrange, Pacific J. Math. 33, 411-428.

Rockafellar, R.T.

1971a) Integrals which are convex functionals II. Pacific J. Math. 39, 429-469.

1971b) Existence and duality theorems for convex problems of Bolza. Trans. Amer. Math. Soc. 159, 1-40.

1972 State constraints in convex problems of Bolza. SIAM J. Contr. 10, 691-715.

1973 Optimal arcs and the minimum value function in problems of Lagrange. Trans. Amer. Math. Soc. 180, 53-83.

1974a) Conjugate Duality and Optimization, No. 16, in Conference Board of Math. Sciences Series. SIAM Publications.

1974b) Convex algebra and duality in dynamic models of produnction. In Mathematical Models in Economics, , Ed., North-Holland.

1975 Existence theorems for general control problems of Bolza and Lagrange. Advances in Math. 15, 312-333.

1976a) Lagrange multipliers in optimization. SIAM-AMS Proceedings, Vol. 9 (R.W. Cottle and C.E. Lemke, eds.), 145-168.

1976b) Integral functionals, normal integrands and measurable selections. Nonlinear Operators and the Calculus of Variations (L. Waelbroeck, ed.), Lecture Notes in Math. No. 543, Springer-Verlag, 157-207.

1979a) Clarke's tangent cones and the boundaries of closed sets in R^n", Nonlin. Analysis Th. Meth. Appl. 3, 145-154.

1979b) Directionally Lipschitzian functions and subdifferential calculus. Proc. London Math. Soc. 39, 331-335.

1979c) La Théorie des Sous-Gradients et Ses Applications à l'Optimisation: Fonctions Convexes et Non Convexes, Collection Chaire Aisenstadt, Presses de l'Université de Montréal, 130 pp. (English translation: The Theory of Subgradients and its Applications to Problems of Optimisation: Convex and Nonconvex Functions, Helderman Verlag, W. Berlin, 1981).

1980 Generalized directional derivatives and subgradients of nonconvex functions. Canadian J. Math. 32, 157-180.

1982a) Proximal subgradients, marginal values, and augmented Lagrangians in nonconvex programming. Math. Op. Res. 6, 427-437.

Rockafellar, R.T.

1982b) Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming. Math. Progr. Study 17, 28-66.

1982c) Favorable classes of Lipschitz continuous functions in subgradient optimization. Nondifferentiable Optimization, E. Nurminski (ed.), Pergamon Press.

1982d) Augmented Lagrangians and marginal values in parametric optimization problems. Generalized Lagrangian Methods in Optimization, A. Wierzbicki (ed.), Pergamon Press.

T.A.a) Marginal values and second-order conditions for optimality. Math. Progr., to appear.

b) Directional differentiability of the optimal value function in a nonlinear programming problem. Math. Progr. Studies, to appear.

Rockafellar, R.T. and R. Wets

1976 Stochastic convex programming. SIAM J. Contr. Opt. 14, 574-589.

T.A. Extended real analysis.

Sargent, R.W.

T.A. On the parametric variation of constraint sets and solutions of minimization problems.

Shi Shu Chung

1980 Remarques sur le gradient généralisé. C.R.A.S. 291, 493-446.

Schinas, J. and M. Boudourides

1981 Higher order differentiability of multifunctions. Nonlinear Analysis. TAM, 5, 509-516.

Strodiot, J.J. and V. Nguyen

1979 Characterisation des solutions optimales en programmation non differentiable. C.R.A.S. 288, 1075-1078.

Sweetser, T.H.

1977 A minimal set-valued strong derivative for vector-valued Lipschitz functions. J. Opt. Th. Appl. 23, 549-562.

Thibault, L.

1979 Cônes tangents et épi-différentiels de fonctions vectorielles. Travaux Sem. Anal. Convexe 9, No. 2, Exp. No. 13.

1982 Subdifferentials of nonconvex vector-valued functions, J. Math. Anal. Appl. 86, 319-344.

T.A.a) On generalized differentials and subdifferentials of Lipschitz vector-valued functions.

b) Tangent cones and quasi-interior tangent cones to multifunctions.

Ursescu, C.

1973 Sur une généralisation de la notion de différentiabilité. Accad. Naz. Lincei. 8, 199-204.

1975 Multifunctions with closed convex graph. Czechoslovak Math. J. 25, 438-441.

Warga, J.

1962 Relaxed variational problems. J. Math. Anal. Appl. 4, 111-128.

1972 Optimal Control of Differential and Functional Equations, Academic Press, New York.

1976 Derivate containers, inverse functions and controllability, in Calculus of Variations and Control Theory, D.L. Russell (ed.), Academic Press, 13-46.

1978a) Controllability and a multiplier rule for nondifferentiable optimization problems. SIAM J. Contr. Opt. 16, 803-812.

1978b) An implicit function theorem without differentiability. Proc. Amer. Math. Soc. 69, 65-69.

Watkins, G.G.

T.A. Clarke's tangent vectors as tangents to Lipschitz continuous curves.

Wolfe, P. and M. Balinski (eds.)

1975 Nondifferentiable optimization. Math. Progr. Studies. North-Holland.

Zowe, J.

1975 A duality theorem for a convex programming theorem in ordered complete vector lattices. J. Math. Anal. Appl. 50, 273-287.

Zlobec, C. and H. Massam

1974 Various definitions of the derivatives in mathematical programming. Math. Progr. 7, 144-161 (Addendum 14, 108-111, 1978).

APPENDIX

Summary of the Calculus of Contingent and Tangent Cones, of Contingent Derivatives and Derivatives of Set-Valued Maps and of Epi-Contingent Derivatives and Epi-Derivatives of Real-Valued Functions.

Derivative of a set-valued map F from X to Y at (x_0,y_0) graph (F)

Operations	derivative	Contingent derivative	derivative of F with closed convex graph
B is a differentiable map from Y to Z $\boxed{BF}$		$D(BF)(x_0,By_0) \supset \nabla B(y_0)DF(x_0,y_0)$	If $B \in L(Y,Z)$, $B.DF(x_0,y_0) = D(BF)(x_0,By_0)$
A is a continuously differentiable map from X_0 to X $\boxed{FA}$	If $X = \mathrm{Im}\nabla A(x_0) - \mathrm{Dom}\, CF(Ax_0,y_0)$ then $C(FA)(x_0,y_0) \supset CF(Ax_0,y_0)\nabla A(x_0)$ and $C(FA)(x_0,y_0)^* \subset \nabla A(x_0)^* CF(Ax_0,y_0)^*$	$D(FA)(x_0,y_0) \subset DF(Ax_0,y_0)\nabla A(x_0)$	If $A \in L(X,Y)$ and $0 \in \mathrm{Int}(\mathrm{Im}\, A - \mathrm{Dom}\, F)$, $D(FA)(x_0,y_0) = DF(Ax_0,y_0)A$ and $D(FA)(x_0,y_0)^* = A^* DF(Ax_0,y_0)$
Let $K \subset X$ be a closed subset and $F\|_K$ denote the restriction of F to K $\boxed{F\|_K}$	If $X = \mathrm{Dom}\, CF(x_0,y_0) - C_K(x_0)$ then $CF(x_0,y_0)\|_{C_K(x_0)} \subset C(F\|_K)(x_0,y_0)$ and $C(F\|_K)(x_0,y_0)^*(\cdot) \subset CF(x_0,y_0)^*(\cdot) + N_K(x_0)$	$DF(x_0,y_0)\|_{T_K(x_0)} \supset D(F\|_K)(x_0,y_0)$	If $0 \in \mathrm{Int}(\mathrm{Dom}\, F - K)$ and K is convex, $DF(x_0,y_0)\|_{T_K(x_0)} = D(F\|_K)(x_0,y_0)$ and $D(F\|_K)(x_0,y_0)^*(0) = DF(x_0,y_0)^*(\cdot) + N_K(x_0)$

Operations	Tangent cone	Contingent cone	Tangent cone to a convex subset
$x \in K \subset L$		$T_K(x) \subset T_L(x)$	$T_K(x) \subset T_L(x)$
$K = \bigcup_{i=1}^{n} K_i$		$T_K(x) = \bigcup_{i=1}^{n} T_{K_i}(x)$	
$K = K_1 \cap K_2$	If $C_{K_1}(x) - C_{K_2}(x) = X$, $C_{K_1}(x) \cap C_{K_2}(x) \subset C_K(x)$	$T_{K_1}(x) \cap T_{K_2}(x) \supset T_K(x)$	If $0 \in \mathrm{Int}(K_1 - K_2)$, $T_{K_1}(x) \cap T_{K_2}(x) = T_K(x)$
$K = \prod_{i=1}^{n} K_i$	$\prod_{i=1}^{n} C_{K_i}(x_i) = C_K(x)$	$\prod_{i=1}^{n} T_{K_i}(x_i) \supset T_K(x)$	$\prod_{i=1}^{n} T_{K_i}(x_i) = T_K(x)$
A is a C^1 - map from X to Y and $K \subset X$.		$\nabla A(x) T_K(x) \subset T_{A(K)}(Ax)$	If $A \in L(X,Y)$, $(AT_K(x)) = T_{A(K)}(Ax)$
A is a C^1 - map from X to Y and $L \subset Y$.	If Im $\nabla A(x) - C_L(Ax) = Y$, $\nabla A(x)^{-1} C_L(Ax) \subset C_{A^{-1}(L)}(x)$	$\nabla A(x)^{-1} T_L(Ax) \supset T_{A^{-1}(L)}(x)$	If $A \in L(X,Y)$ and $0 \in \mathrm{Int}(\mathrm{Im}\, A - L)$, $A^{-1} T_{L(Ax)} = T_{A^{-1}(L)}(x)$

Epi-derivative of a proper real-valued function $V : X \to \mathbb{R} \cup \{+\infty\}$ at $x_0 \in \mathrm{Dom}\, V$			
Operations	epi-derivative of a lower semicontinuous function when $\dim(X) < +\infty$	epi-contingent derivative	epi-derivative of a lower semicontinuous convex function
Let A be a continuously differentiable map from X_0 to X VA	If $X = \mathrm{Im}\nabla A(x_0) - \mathrm{Dom}\, C_+V(Ax_0)$, then $C_+V(Ax_0)\nabla A(x_0) \geq C_+(VA)(x_0)$ and $\partial V(Ax_0) \subset \nabla A(x_0)^*\partial V(x_0)$	$D_+V(Ax_0)\nabla A(x_0) \leq D_+(VA)(x_0)$	If $A \in \mathcal{L}(X_0,X)$ and $0 \in \mathrm{Int}(\mathrm{Im}A - \mathrm{Dom}\, V)$, then $D_+V(Ax_0)(\nabla A(x_0)) = D_+(VA)(x_0)$ and $\partial(VA)(x_0) = A^*\partial V(Ax_0)$
Let K be a closed subset of X $V\|_K$	If $X = C_K(x_0) - \mathrm{Dom}\, C_+V(x_0)$ then $C_+(V\|_K)(x_0) \leq C_+V(x_0)\|_{C_K(x_0)}$ and $\partial(V\|_K)(x_0) \subset \partial V(x_0) + N_K(x_0)$	$D_+(V\|_K)(x_0) \geq D_+V(x_0)\|_{T_K(x_0)}$	$D_+(V\|_K)(x_0) = D_+V(x_0)\|_{T_K(x_0)}$ and $\partial(V\|_K)(x_0) = \partial V(x_0) + N_K(x_0)$ when K is convex and $0 \in \mathrm{Int}(K - \mathrm{Dom}\, V)$
Let $V_i : X \to \mathbb{R} \cup \{+\infty\}$ $\max V_i$ If $x_0 \in \bigcap_{i=1}^{n} \mathrm{Int}\, \mathrm{Dom}\, V_i$, we set $J(x_0) = \{ i \mid \max_i V_i(x_0) = V_j(x_0)\}$	If the functions are locally lipschitz, $C_+(\max V_i)(x_0) \leq \max_{j \in J(x_0)} C_+V_j(x_0)$ and $\partial(\max V_i)(x_0) \subset \overline{co} \bigcup_{j \in J(x_0)} \partial V_j(x_0)$	$D_+(\max V_i)(x_0) \geq \max_{j \in J(x_0)} D_+V_j(x_0)$	$D_+(\max V_i)(x_0) = \max_{j \in J(x_0)} D_+V_j(x_0)$ and $\partial(\max V_i)(x_0) = \overline{co} \bigcup_{j \in J(x_0)} \partial V_j(x_0)$
V and W are proper functions and $x_0 \in \mathrm{Dom}\, V \cap \mathrm{Dom}\, W$ $V + W$	If V is lipschitz around x_0, then $C_+(V+W)(x_0) \leq C_+V(x_0) + C_+W(x_0)$ and $\partial(V+W)(x_0) \subset \partial V(x_0) + \partial W(x_0)$	$D_+(V+W)(x_0) \geq D_+V(x_0) + D_+W(x_0)$	If $0 \in \mathrm{Int}(\mathrm{Dom}\, V - \mathrm{Dom}\, W)$ then $D_+(V+W)(x_0) = D_+V(x_0) + D_+W(x_0)$ and $\partial(V+W)(x_0) = \partial V(x_0) + \partial W(x_0)$